

Université de Montréal

Estimation utilisant les polynômes de Bernstein

par

Tchouake Tchuiguetp Hervé

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Statistique

mars 2013

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

Estimation utilisant les polynômes de Bernstein

présenté par

Tchouake Tchuiguetp Hervé

a été évalué par un jury composé des personnes suivantes :

Pierre Lafaye De Micheaux

(président-rapporteur)

François Perron

(directeur de recherche)

Louis-Pierre Arguin

(membre du jury)

Mémoire accepté le:

3 mars 2013

SOMMAIRE

Ce mémoire porte sur la présentation des estimateurs de Bernstein qui sont des alternatives récentes aux différents estimateurs classiques des fonctions de répartition et de densité. Plus précisément, nous étudions leurs différentes propriétés et les comparons à celles de la fonction de répartition empirique et à celles de l'estimateur obtenu par la méthode du noyau. Nous déterminons une expression asymptotique des deux premiers moments de l'estimateur de Bernstein pour la fonction de répartition. Comme pour les estimateurs classiques, nous montrons que cet estimateur vérifie la propriété de Chung-Smirnov sous certaines conditions. Nous montrons ensuite que l'estimateur de Bernstein est meilleur que la fonction de répartition empirique en terme d'erreur quadratique moyenne. En s'intéressant au comportement asymptotique des estimateurs de Bernstein, pour un choix convenable du degré du polynôme, nous montrons que ces estimateurs sont asymptotiquement normaux. Des études numériques sur quelques distributions classiques nous permettent de confirmer que les estimateurs de Bernstein peuvent être préférables aux estimateurs classiques.

Mots clés : Estimation non paramétrique, propriétés asymptotiques, processus empiriques, convergence presque sûre, étude par simulation.

SUMMARY

This thesis focuses on the presentation of the Bernstein estimators which are recent alternatives to conventional estimators of the distribution function and density. More precisely, we study their various properties and compare them with the empirical distribution function and the kernel method estimators. We determine an asymptotic expression of the first two moments of the Bernstein estimator for the distribution function. As the conventional estimators, we show that this estimator satisfies the Chung-Smirnov property under conditions. We then show that the Bernstein estimator is better than the empirical distribution function in terms of mean squared error. We are interested in the asymptotic behavior of Bernstein estimators, for a suitable choice of the degree of the polynomial, we show that the Bernstein estimators are asymptotically normal. Numerical studies on some classical distributions confirm that the Bernstein estimators may be preferable to conventional estimators.

Keywords : Non-parametric density estimator ; asymptotic properties ; empirical processes ; almost sure limits ; simulation study.

TABLE DES MATIÈRES

Sommaire	v
Summary	vii
Liste des figures	xi
Remerciements	1
Introduction	3
Chapitre 1. Polynômes de Bernstein	5
Chapitre 2. Processus empiriques	13
2.1. Inégalités importantes.....	15
2.2. Inégalités de Bennett-Wellner.....	19
2.3. Preuve du Théorème 2.1	23
Chapitre 3. Estimateurs de Bernstein	29
3.1. La fonction de répartition empirique	29
3.2. Estimateurs de Bernstein	32
3.2.1. Définitions et Propriétés	32
3.2.2. Loi du logarithme itéré	44
3.2.3. Résultats asymptotiques	48
Chapitre 4. Etudes numériques	55
4.1. Fonction de répartition.....	55

4.2. Densité	64
Conclusion	71
Bibliographie	73
Annexe A. Résultats importants	A-i
A.1. Théorème de Lindeberg	A-i
A.2. Théorème de Dini	A-ii
Annexe B. Codes R utilisés pour les simulations	B-i
B.1. Cas (iii) section 4.1	B-i
B.2. Cas (iii) section 4.2	B-iv

LISTE DES FIGURES

4.1	Estimateur de Bernstein pour un échantillon de loi $\text{Beta}(3,2)$	57
4.2	Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(1,6)$	58
4.3	Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$...	59
4.4	Erreur d'estimation pour un échantillon de loi $\text{Beta}(3,2)$	61
4.5	Valeur optimale de m pour un échantillon de loi $\text{Beta}(3,2)$	61
4.6	Erreur d'estimation pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$	62
4.7	Valeur optimale de m pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$	62
4.8	Erreur d'estimation pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$	63
4.9	Valeur optimale de m pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$	63
4.10	Estimateur de Bernstein pour un échantillon de loi $\text{Beta}(3,2)$: cas m optimal.	65
4.11	Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$: cas m optimal.	66
4.12	Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$: cas m optimal.	67
4.13	Estimateur de Bernstein pour un échantillon de loi $\text{Beta}(2,5)$	68
4.14	Estimateur de Bernstein pour un échantillon de loi $\chi^2(2)$	69
4.15	Estimateur de Bernstein pour un échantillon de loi $t(2)$	70

REMERCIEMENTS

Mes remerciements vont premièrement à mon directeur de recherche ; le professeur François Perron, l'aide et le support qu'il m'a apportés ont grandement favorisé mon épanouissement intellectuel durant ma maîtrise au département de mathématiques et de statistique de l'Université de Montréal.

J'aimerais également remercier tout le personnel enseignant et non-enseignant du département de mathématiques et statistique pour la grande qualité de la formation et du soutien dont j'ai bénéficié durant ma maîtrise.

Je remercie tous mes camarades et amis qui ont rendu ce travail plus agréable : David, Fabiola, Hervé, Thanina, Joseph, Joël, Thierry et tous ceux qui de près ou de loin m'ont aidé lors de la rédaction de ce travail.

Je remercie enfin ma famille pour ses conseils et son précieux support.

INTRODUCTION

Plusieurs phénomènes de la vie courante, comme la durée de vie d'une ampoule électrique, l'âge des participants à une étude ou le gain d'un joueur dans un jeu de hasard peuvent être modélisés par des variables aléatoires. Partant de l'observation d'un tel phénomène, on aimerait avoir une idée précise de son ampleur. Les fonctions de répartition et de densité sont des concepts essentiels en statistique. Ils permettent de décrire et de caractériser la distribution d'une variable aléatoire. Si on dispose d'un ensemble de mesures observées formant un échantillon d'une loi, qui en pratique est souvent inconnue, on aimerait construire à partir de cet échantillon un estimateur pouvant approcher avec une certaine précision la fonction de densité (ou la fonction de répartition). Une approche simple est la méthode paramétrique. Elle suppose que les données proviennent d'une loi paramétrique connue. Il suffit alors dans ce cas d'estimer les paramètres caractérisant cette loi pour déduire une estimation de sa densité. Nous allons nous intéresser dans ce mémoire à l'approche non paramétrique. Un estimateur non paramétrique intuitif est l'histogramme qui permet de dégager l'allure de la fonction de densité. La fonction de répartition empirique est connue pour être un bon estimateur de la fonction de répartition. Moyennant des informations a priori sur la distribution de l'échantillon, on peut construire des estimateurs s'ajustant mieux à la nature de la variable d'intérêt. Nous nous intéresserons dans ce travail aux estimateurs basés sur les polynômes de Bernstein. Ces polynômes furent introduits par Bernstein (1912), dans le but de donner une preuve du théorème de Weierstrass. En s'inspirant de ce résultat d'approximation uniforme, Vitale (1975) définit l'estimateur de Bernstein pour la densité et étudia ses premières propriétés. Plus récemment, Babu et al. (2002) proposèrent un estimateur de la fonction de répartition basé sur les polynômes de Bernstein. En s'appuyant sur des résultats de

Einmahl (1987) sur les processus empiriques, et de Lorentz (1986) sur les relations entre le module de continuité et les polynômes de Bernstein, Leblanc (2009, 2012a) montra que les estimateurs de Bernstein possèdent sous certaines conditions des propriétés similaires à celles des estimateurs classiques comme la fonction de répartition empirique et les estimateurs de la méthode des noyaux.

Ce mémoire sera structuré de la manière suivante. Au chapitre 1, nous présenterons les résultats fondamentaux concernant les polynômes de Bernstein. Ces résultats concernent le comportement asymptotique des polynômes de Bernstein et leur lien avec le module de continuité. Au chapitre 2, nous nous intéresserons aux processus empiriques multivariés, nous présenterons plusieurs inégalités très utiles, dont celles de Bennett-Wellner pour le module de continuité du processus empirique ; le chapitre 3 sera la partie centrale du travail. Il utilisera les deux premiers chapitres pour donner des résultats importants concernant des estimateurs de Bernstein : théorèmes de convergence et asymptotiques, l'inégalité de Chung-Smirnov et comparaison avec la fonction de répartition empirique. Au chapitre 4, en partant de quelques distributions bien connues, nous présenterons quelques résultats de simulations dans le but de confirmer les résultats du chapitre 3.

Chapitre 1

POLYNÔMES DE BERNSTEIN

Nous définissons ici les polynômes de Bernstein et nous présentons leurs propriétés principales. Les différents résultats sont tirés du chapitre 1 de Lorentz (1986).

Soit u une fonction prenant ses valeurs dans l'intervalle $I = [0, 1]$. Le polynôme de Bernstein d'ordre m associé à u , noté $B_m(u)$, est défini par :

$$B_m(u)(x) = \sum_{k=0}^m u(k/m) b_k(m, x),$$

où

$$b_k(m, x) = \binom{m}{k} x^k (1-x)^{m-k}, \quad k = 0, 1, \dots, m,$$

sont les termes binomiaux.

L'étude des polynômes de Bernstein est donc très liée aux coefficients binomiaux. Nous allons maintenant donner quelques résultats utiles concernant les termes binomiaux $b_k(m, x)$. Considérons les termes :

$$T_s(m, x) = \sum_{k=0}^m (k - mx)^s b_k(m, x), \quad s = 0, 1, 2, \dots$$

On a le lemme suivant.

Lemme 1.1. *Considérons $m \in \mathbb{N}_0 = \mathbb{N} \setminus \{0\}$ et $x \in I$ quelconques.*

(i) *On a les résultats suivants.*

$$T_0(m, x) = 1, \tag{1.1a}$$

$$T_1(m, x) = 0, \tag{1.1b}$$

$$T_2(m, x) = mx(1-x). \tag{1.1c}$$

(ii) Pour m fixé, $T_{2s}(m, x)$ est un polynôme en $X = x(1 - x)$ et $T_{2s+1}(m, x)$ est un polynôme en X multiplié par $(1 - 2x)$, pour tout $s \geq 1$.

(iii) Pour x fixé, $T_s(m, x)$ est un polynôme en m de degré $\lfloor s/2 \rfloor$.

Preuve. Le résultat (1.1a) est immédiat. Soit Y une variable aléatoire de loi binomiale de paramètres m et x . Un calcul simple montre que $E(Y) = \sum_{k=0}^m kb_k(m, x) = mx$ et

alors $T_1(m, x) = \sum_{k=0}^m kb_k(m, x) - mxT_0(m, x) = 0$, d'où (1.1b). On a aussi

$$\text{Var}(Y) = \sum_{k=0}^m (k - mx)^2 b_k(m, x) = T_2(m, x) = mx(1 - x),$$

ce qui montre (1.1c). Par dérivation, on a :

$$\begin{aligned} \frac{d}{dx} T_s(m, x) &= -msT_{s-1}(m, x) + \sum_{k=0}^m (k - mx)^s \frac{d}{dx} b_k(m, x), \\ &= -msT_{s-1}(m, x) + [x(1 - x)]^{-1} \sum_{k=0}^m (k - mx)^{s+1} b_k(m, x). \end{aligned}$$

On déduit la formule de récurrence

$$T_{s+1}(m, x) = x(1 - x) \left[\frac{d}{dx} T_s(m, x) + msT_{s-1}(m, x) \right]. \quad (1.2)$$

On déduit (ii) et (iii) de (1.2) par récurrence sur s . \square

Puisque $T_{2s}(m, x)$ est un polynôme de degré s en m et un polynôme en $x(1 - x)$, il existe une constante $C_s > 0$ telle que $0 \leq T_{2s}(m, x) \leq C_s m^s$ pour tout $x \in I$. On déduit alors que pour tout $\alpha > 0$ et $s \geq 0$,

$$\begin{aligned} \sum_{\substack{k=0 \\ \left| x - \frac{k}{m} \right| \geq \alpha}}^m b_k(m, x) &\leq \alpha^{-2s} \sum_{\substack{k=0 \\ \left| x - \frac{k}{m} \right| \geq \alpha}}^m \left(x - \frac{k}{m} \right)^{2s} b_k(m, x), \\ &\leq \alpha^{-2s} m^{-2s} T_{2s}(m, x) \leq C_s \alpha^{-2s} m^{-s}. \end{aligned} \quad (1.3)$$

Dans toute la suite, si u est une fonction bornée sur I , on note

$$\|u\| = \sup_{x \in I} |u(x)|.$$

Les polynômes de Bernstein furent proposés afin de prouver le théorème de Weierstrass qui s'énonce de la manière suivante : toute fonction continue à valeurs réelles définie sur un intervalle fermé et borné de \mathbb{R} peut être approchée uniformément par une suite de polynômes. Le résultat suivant traite un cas particulier.

Théorème 1.2. *Si u est une fonction continue sur I , alors $B_m(u)(x) \rightarrow u(x)$, lorsque $m \rightarrow \infty$ uniformément sur I , c'est-à-dire :*

$$\|B_m(u) - u\| \rightarrow 0, \quad m \rightarrow \infty. \quad (1.4)$$

Preuve. Soit $x \in I$ et m un entier naturel non nul. On a :

$$\begin{aligned} |u(x) - B_m(u)(x)| &= \left| u(x) - \sum_{k=0}^m u(k/m) b_k(m, x) \right|, \\ &= \left| \sum_{k=0}^m u(x) b_k(m, x) - \sum_{k=0}^m u(k/m) b_k(m, x) \right|, \\ &= \left| \sum_{k=0}^m (u(x) - u(k/m)) b_k(m, x) \right|, \\ &\leq \sum_{k=0}^m |u(x) - u(k/m)| b_k(m, x). \end{aligned}$$

La fonction u étant continue sur le compact I , elle est uniformément continue et est bornée. Il existe donc $M > 0$ tel que $|u(x)| \leq M$ pour tout $x \in I$. Soit maintenant $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout $x, y \in I$; $|x - y| \leq \alpha \Rightarrow |u(x) - u(y)| \leq \frac{\varepsilon}{2}$. Soit $x \in I$, en utilisant (1.1a), (1.1c) et (1.3) on peut alors écrire :

$$\begin{aligned} |u(x) - B_m(u)(x)| &\leq \sum_{\substack{k=0 \\ \left|x - \frac{k}{m}\right| \leq \alpha}}^m |u(x) - u(k/m)| b_k(m, x) + \sum_{\substack{k=0 \\ \left|x - \frac{k}{m}\right| > \alpha}}^m |u(x) - u(k/m)| b_k(m, x), \\ &\leq \frac{\varepsilon}{2} \sum_{\substack{k=0 \\ \left|x - \frac{k}{m}\right| \leq \alpha}}^m b_k(m, x) + 2M \sum_{\substack{k=0 \\ \left|x - \frac{k}{m}\right| > \alpha}}^m b_k(m, x), \\ &\leq \frac{\varepsilon}{2} T_0(m, x) + \frac{2M}{m^2 \alpha^2} T_2(m, x), \\ &\leq \frac{\varepsilon}{2} + \frac{M}{2m\alpha^2}. \end{aligned}$$

Il existe $m_0 \in \mathbb{N}$ tel que $m \geq m_0 \Rightarrow \frac{M}{2m\alpha^2} \leq \frac{\varepsilon}{2}$. On a finalement pour $m \geq m_0$, $|u(x) - B_m(u)(x)| \leq \varepsilon$ quelque soit $x \in I$. \square

Définissons maintenant quelques notions importantes.

Définition 1.1. Soit u une fonction définie sur I . Elle sera dite de Lipschitz d'ordre $\alpha \in (0, 1]$, s'il existe une constante $c > 0$ telle que :

$$|u(x) - u(y)| \leq c|x - y|^\alpha, \quad (1.5)$$

pour tout $x, y \in I$. On notera par $\text{Lip}(\alpha)$ l'ensemble de fonctions vérifiant (1.5).

Définition 1.2. Soit u une fonction bornée sur I . On appelle module de continuité associé à u , l'application ω_u définie sur I par :

$$\omega_u(t) = \sup \left\{ |u(x) - u(y)| ; x, y \in I, |x - y| \leq t \right\}.$$

La fonction ω_u vérifie les propriétés suivantes :

- 1) ω_u est croissante et nulle en 0,
- 2) ω_u est sous additive, c'est-à-dire $\omega_u(t_1 + t_2) \leq \omega_u(t_1) + \omega_u(t_2)$ pour $t_1, t_2 \in I$ tels que $t_1 + t_2 \in I$,
- 3) ω_u est continue si u est uniformément continue,
- 4) $\omega_u(\lambda t) \leq (\lambda + 1)\omega_u(t)$, $t \in I$ et $\lambda > 0$.

Les différentes propriétés du module de continuité se vérifient aisément.

Notons qu'il existe plusieurs définitions équivalentes du module de continuité ; une autre très souvent rencontrée dans la littérature est :

$$\omega_u(t) = \sup_{0 \leq h \leq t} \sup_{0 \leq x \leq 1-h} |u(x+h) - u(x)|.$$

Nous allons maintenant présenter quelques résultats concernant le module de continuité.

Théorème 1.3. Soient u une fonction continue et $\omega_u \equiv \omega$ son module de continuité. On a :

$$|u(x) - B_m(u)(x)| \leq \frac{3}{2}\omega(m^{-1/2}). \quad (1.6)$$

Preuve. Pour $x_1, x_2 \in I$ et $\delta > 0$, définissons l'entier $\lambda \equiv \lambda(x_1, x_2; \delta) = \lfloor |x_1 - x_2|\delta^{-1} \rfloor$. On peut écrire la différence $u(x_1) - u(x_2)$ comme une somme de $\lambda + 1$ différences de la forme $u(x) - u(x')$ avec $|x - x'| \leq \delta$. On déduit alors que :

$$|u(x_1) - u(x_2)| \leq (\lambda + 1)\omega(\delta). \quad (1.7)$$

Une application de l'inégalité de Cauchy-Schwarz puis l'utilisation des relations (1.1a) et (1.1c) donne :

$$\begin{aligned}
 \sum_{k=0}^m |\chi - k/m| b_k(m, \chi) &= \sum_{k=0}^m b_k(m, \chi)^{1/2} |\chi - k/m| b_k(m, \chi)^{1/2}, \\
 &\leq \left(\sum_{k=0}^m (\chi - k/m)^2 b_k(m, \chi) \right)^{1/2} \left(\sum_{k=0}^m b_k(m, \chi) \right)^{1/2}, \\
 &\leq (4m)^{-1/2}.
 \end{aligned} \tag{1.8}$$

En utilisant (1.7) et (1.8), on peut alors écrire :

$$\begin{aligned}
 |u(\chi) - B_m(u)(\chi)| &\leq \sum_{k=0}^m |u(\chi) - u(k/m)| b_k(m, \chi), \\
 &\leq \omega(\delta) \sum_{k=0}^m (1 + \lambda(\chi, k/m; \delta)) b_k(m, \chi), \\
 &= \omega(\delta) \left(1 + \sum_{k=0}^m \lambda(\chi, k/m; \delta) b_k(m, \chi) \right), \\
 &\leq \omega(\delta) \left(1 + \delta^{-1} \sum_{k=0}^m |\chi - k/m| b_k(m, \chi) \right), \\
 &\leq \omega(\delta) (1 + (4m\delta^2)^{-1/2}).
 \end{aligned}$$

Il suffit que prendre $\delta = m^{-1/2}$ pour conclure. \square

Théorème 1.4. Soient u une fonction continûment dérivable et ω_1 le module de continuité de u' . on a :

$$|u(\chi) - B_m(u)(\chi)| \leq \frac{3}{4} m^{-1/2} \omega_1(m^{-1/2}). \tag{1.9}$$

Preuve. Soient $x_1, x_2 \in I$ vérifiant $x_1 < x_2$. D'après le théorème des accroissements finis, il existe $\xi \in (x_1, x_2)$ tel que :

$$u(x_1) - u(x_2) = (x_1 - x_2)u'(\xi) = (x_1 - x_2)u'(x_1) + (x_1 - x_2)(u'(\xi) - u'(x_1)).$$

En raisonnant comme dans la preuve du Théorème 1.3 et en utilisant (1.1b), on a :

$$\begin{aligned}
|u(x) - B_m(u)(x)| &= \left| \sum_{k=0}^m (u(x) - u(k/m)) b_k(m, x) \right|, \\
&\leq \left| \sum_{k=0}^m (x - k/m) u'(x) b_k(m, x) \right| + \omega_1(\delta) \sum_{k=0}^m (1 + \lambda) |x - k/m| b_k(m, x), \\
&\leq \omega_1(\delta) \left(\sum_{k=0}^m |x - k/m| b_k(m, x) + \sum_{k=0}^m |x - k/m| \lambda(x, k/m; \delta) b_k(m, x) \right), \\
&\leq \omega_1(\delta) \left(\sum_{k=0}^m |x - k/m| b_k(m, x) + \delta^{-1} \sum_{k=0}^m (x - k/m)^2 b_k(m, x) \right).
\end{aligned}$$

D'après (1.1c), le second terme du membre de droite de l'inégalité ci-dessus se majore par $\frac{1}{4}(m\delta)^{-1}$. On déduit alors que :

$$|u(x) - B_m(u)(x)| \leq \omega_1(\delta) \left(\frac{1}{2} m^{-1/2} + \frac{1}{4} (m\delta)^{-1} \right).$$

Il suffit alors de prendre $\delta = m^{-1/2}$ pour avoir le résultat. \square

Le lemme suivant donne un résultat très utile dans la suite.

Lemme 1.5. *Si u est deux fois continûment dérivable à dérivées bornées, alors pour $x \in I$ on a :*

$$\lim_{m \rightarrow \infty} m(B_m(u)(x) - u(x)) = x(1-x)u''(x)/2.$$

Preuve. D'après la formule de Taylor, on peut écrire

$$u(k/m) = u(x) + (k/m - x)u'(x) + (k/m - x)^2 [u''(x)/2 + \eta(k/m - x)], \quad (1.10)$$

où η est une fonction bornée vérifiant $\lim_{h \rightarrow 0} \eta(h) = 0$. Il existe une constante H telle $|\eta(h)| \leq H$ pour tout h . On multiplie les deux membres (1.10) par $b_k(m, x)$, en sommant sur $0 \leq k \leq m$ on obtient alors :

$$B_m(u)(x) = u(x) + m^{-2} u''(x) T_2(m, x)/2 + \sum_{k=0}^m (x - k/m)^2 \eta(k/m - x) b_k(m, x).$$

Soit $\varepsilon > 0$. Il existe $\delta > 0$ tel que $|h| \leq \delta$ implique $|\eta(h)| \leq \varepsilon$. En utilisant (1.3), on a alors :

$$\begin{aligned}
\sum_{k=0}^m (x - k/m)^2 \eta(k/m - x) b_k(m, x) &\leq \varepsilon \sum_{\substack{k=0 \\ |x - k/m| \leq \delta}}^m (x - k/m)^2 b_k(m, x) + H \sum_{\substack{k=0 \\ |x - k/m| > \delta}}^m b_k(m, x), \\
&\leq \varepsilon m^{-2} T_2(m, x) + CHm^{-2}.
\end{aligned}$$

Par ailleurs, il existe m_0 tel que $m \geq m_0$ implique $CHm^{-1} \leq \varepsilon$. On déduit alors d'après (1.1c) que pour $m \geq m_0$,

$$B_m(u)(x) = u(x) + \frac{x(1-x)}{2m} u''(x) + \frac{\varepsilon_m}{m}, \quad (1.11)$$

avec $\varepsilon_m \rightarrow 0$, $m \rightarrow \infty$. Et le lemme en découle. \square

Chapitre 2

PROCESSUS EMPIRIQUES

Les processus empiriques sont d'une importance majeure par leurs applications dans les sciences appliquées comme la physique et l'ingénierie. Dans ce chapitre, nous allons nous intéresser au cas multivarié en partant du travail de Einmahl (1987).

Soit $X_1, X_2, \dots, X_n, \dots$ une suite de vecteurs aléatoires i.i.d définis sur un espace probabilisé (Ω, \mathcal{F}, P) à valeurs dans I^d , $I = [0, 1]$, $d \in \mathbb{N}_0 = \mathbb{N} \setminus \{0\}$. Soit F la fonction de répartition commune des vecteurs X_i .

On désigne par \mathcal{R} l'ensemble des pavés de la forme

$$R(t, s) = [s_1, t_1] \times [s_2, t_2] \times \dots \times [s_d, t_d],$$

et posons $|R(t, s)| = |t - s| = \prod_{i=1}^d (t_i - s_i)$, $s \leq t$, $s, t \in I^d$.

La fonction de distribution empirique est définie pour $t \in I^d$ par :

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq t\}}. \quad (2.1)$$

Pour $R \in \mathcal{R}$, on définit

$$F(R) = P\{X_i \in R\}, \quad F_n(R) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \in R\}}.$$

Le processus empirique multivarié est défini par :

$$U_n^F(t) = \sqrt{n}(F_n(t) - F(t)).$$

Pour $R \in \mathcal{R}$, on définit

$$U_n^F(R) = \sqrt{n}(F_n(R) - F(R)).$$

Le module de continuité du processus \mathbb{U}_n^F , appelé module d'oscillation sera noté

$$\omega_n^F(\alpha) = \sup_{|R| \leq \alpha} |\mathbb{U}_n^F(R)|, \quad 0 \leq \alpha \leq 1. \quad (2.2)$$

On rencontre aussi dans la littérature la définition suivante :

$$M_n^F(\alpha) = \sup_{|t_j - s_j| \leq \alpha, \forall j} |\mathbb{U}_n^F(t) - \mathbb{U}_n^F(s)|, \quad \alpha \geq 0. \quad (2.3)$$

Notons que les définitions (2.2) et (2.3) du module d'oscillation coïncident dans le cas univarié. Nous noterons $\mathbb{U}_n^F \equiv \mathbb{U}_n$, $\omega_n^F \equiv \omega_n$ et $M_n^F \equiv M_n$ si F est de loi uniforme sur I^d .

Le but de ce chapitre est de démontrer le théorème suivant :

Théorème 2.1. (Voir Einmahl (1987)). Soit $\{a_n\}$ ($a_n \in (0, 1)$) une suite vérifiant $a_n \searrow 0$ quand $n \rightarrow \infty$.

(a) Si $\log(1/a_n)/\log \log n \rightarrow c \in [0, \infty)$ et $na_n \nearrow$ quand $n \rightarrow \infty$, alors

$$\limsup_{n \rightarrow \infty} \frac{\omega_n(a_n)}{(a_n \log \log n)^{1/2}} = (2(1+c))^{1/2} \text{ p.s.}$$

(b) Si $\log(1/a_n)/\log \log n \rightarrow \infty$, $na_n/\log n \rightarrow \infty$ et $na_n \nearrow$ quand $n \rightarrow \infty$, alors

$$\lim_{n \rightarrow \infty} \frac{\omega_n(a_n)}{(a_n \log(1/a_n))^{1/2}} = 2^{1/2} \text{ p.s.}$$

(c) Si $na_n/\log n \rightarrow c \in (0, \infty)$ quand $n \rightarrow \infty$, alors

$$\lim_{n \rightarrow \infty} \frac{n^{1/2} \omega_n(a_n)}{\log n} = c(\beta_c - 1) \text{ p.s.}$$

où β_c est défini par $h(\beta_c) \equiv \beta_c(\log \beta_c - 1) + 1 = 1/c$ et $\beta_c > 1$.

(d) Si $na_n/\log n \rightarrow 0$, $\log(1/a_n)/\log n \rightarrow 1$ et $n^2 a_n \nearrow$ quand $n \rightarrow \infty$, alors

$$\lim_{n \rightarrow \infty} \frac{n^{1/2} \omega_n(a_n) \log((\log n)/(na_n))}{\log n} = 1 \text{ p.s.}$$

(e) Si $\log(1/a_n)/\log n = c \in (1, \infty)$, (i.e $a_n = n^{-c}$), alors

$$\limsup_{n \rightarrow \infty} n^{1/2} \omega_n(a_n) = c/(c-1) \text{ p.s.,}$$

si de plus $c/(c-1) \notin \mathbb{N}$, on peut remplacer c par $c_n \rightarrow c$ et \limsup par \lim .

Commençons par prouver quelques inégalités importantes.

2.1. INÉGALITÉS IMPORTANTES

Pour s, v et $u \in \mathbb{N}^d$ vérifiant $s \leq u \leq v$, soit $\{\xi_t\}_{t \leq v}$ une suite de variables aléatoires indépendantes, de moyenne nulle et de variance $\sigma_t^2 < \infty$.

Posons

$$S_v = \sum_{t \leq v} \xi_t, \quad S_{s,u} = \sum_{s \leq t \leq u} \xi_t.$$

et $\sigma^2 = \text{Var}(S_v) = \sum_{t \leq v} \sigma_t^2$,

On a l'inégalité suivante :

Proposition 2.2. *Pour tout $\lambda \in \mathbb{R}$, on a :*

$$P\left\{\max_{u \leq v} S_u \geq \lambda\right\} \leq 2^d P\left\{S_v \geq \lambda - d\sqrt{2\sigma^2}\right\}. \quad (2.4)$$

Preuve. Posons $A = \left\{\max_{u \leq v} S_u \geq \lambda\right\}$ et $B = \left\{\max_{u \leq v, u_d = v_d} S_u \geq \lambda - \sqrt{2\sigma^2}\right\}$.
Commençons par montrer que :

$$P(A) \leq 2P(B).$$

Soient $C = \left\{k \in \mathbb{N}_0 : k \leq \prod_{j=1}^d v_j\right\}$ et $D = \left\{t \in \mathbb{N}_0^d : t \leq v\right\}$. On a $\#C = \#D = \prod_{j=1}^d v_j$.
On peut construire une bijection $g : C \rightarrow D$ telle que : $k < l \Rightarrow g(k)_d \leq g(l)_d$. Posons $\tilde{g}(k) = (g(k)_1, \dots, g(k)_{d-1}, v_d)$, $E_k = \{S_{g(k)} \geq \lambda\}$ et $A_k = E_k \cap E_{k-1}^c \cap E_{k-2}^c \cap \dots \cap E_1^c$.
Clairement $(A_k)_{k \in C}$ forme une partition de A . Remarquons que

$$S_{\tilde{g}(k)} - S_{g(k)} = \sum_{t \in F_k} \xi_t,$$

où $F_k = \{t \in \mathbb{N}_0^d : t_1 \leq g(k)_1, \dots, t_{d-1} \leq g(k)_{d-1}, g(k)_d < t_d \leq v_d\}$.

Puisque $k < l \Rightarrow g(k)_d \leq g(l)_d$, aucun multi-indice de F_k n'est présent dans les sommes définissant l'évènement A_k . On en déduit que A_k et $\{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}$ sont indépendants. L'application de l'inégalité de Tchebychev à la variable aléatoire $S_{g(k)} - S_{\tilde{g}(k)}$ donne :

$$P\{S_{g(k)} - S_{\tilde{g}(k)} \geq \sqrt{2\sigma^2}\} \leq \frac{\text{Var}(S_{g(k)} - S_{\tilde{g}(k)})}{2\sigma^2} \leq \frac{\sigma^2}{2\sigma^2} = \frac{1}{2},$$

on déduit alors par passage au complémentaire que :

$$\frac{1}{2} \leq P\{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}.$$

Et alors par indépendance

$$\begin{aligned}
P(A_k) &\leq 2P(A_k)P\{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}, \\
&= 2P\left\{A_k \cap \{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}\right\}, \\
&\leq 2P(A_k \cap B).
\end{aligned}$$

En effet $\{S_{g(k)} \geq \lambda\}$ et $\{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}$, impliquent $\{S_{\tilde{g}(k)} \geq \lambda - \sqrt{2\sigma^2}\}$.

On déduit alors que :

$$P(A) = \sum_{k \in C} P(A_k) \leq 2 \sum_{k \in C} P(A_k \cap B) = 2P(A \cap B) \leq 2P(B).$$

On obtient alors le (2.4) par un raisonnement inductif. \square

Proposition 2.3. *Pour tout $\lambda \in \mathbb{R}$, on a :*

$$P\left\{\max_{s \leq u \leq v} S_{s,u} \geq \lambda\right\} \leq 2^{2d} P\left\{S_v \geq \lambda - 2d\sqrt{2\sigma^2}\right\}. \quad (2.5)$$

Preuve. On va construire une suite de variables aléatoires indexées par des éléments de \mathbb{N}^{2d} , vérifiant les hypothèses d'application de la Proposition 2.2. Notons par \bar{s} et \bar{t} les éléments de \mathbb{N}^{2d} . Soit $\bar{t} \leq \bar{w} = (v_1, \dots, v_d, v_1, \dots, v_d)$, vérifiant $\bar{t}_j + \bar{t}_{d+j} = v_j + 1$ pour tout $1 \leq j \leq d$. Définissons $\xi_{\bar{t}} = \xi_{(\bar{t}_1, \dots, \bar{t}_d)}$ et $\xi_{\bar{t}} = 0$ dans les autres cas.

Il est facile de voir que :

$$S_{s,u} = S_{(u_1, \dots, u_d, v_1+1-s_1, \dots, v_d+1-s_d)},$$

d'après (2.4), on peut ainsi écrire

$$\begin{aligned}
P\left\{\max_{s \leq u \leq v} S_{s,u} \geq \lambda\right\} &\leq P\left\{\max_{\bar{u} \leq \bar{w}} S_{\bar{u}} \geq \lambda\right\}, \\
&\leq 2^{2d} P\left\{S_{\bar{w}} \geq \lambda - 2d\sqrt{2\sigma^2}\right\} = 2^{2d} P\left\{S_v \geq \lambda - 2d\sqrt{2\sigma^2}\right\}.
\end{aligned}$$

\square

Nous allons maintenant définir le processus de Poisson multivarié.

Définition 2.1. Un processus de comptage multivarié est un processus $\{N(t), t \geq 0\}$ vérifiant les propriétés suivantes :

- (1) $N(t) \geq 0$.
- (2) $N(t)$ est à valeurs entières et croissant.

Définition 2.2. Un processus de Poisson non homogène d'intensité positive $\lambda(t)$ est un processus de comptage multivarié $\{N(t), t \geq 0\}$ vérifiant les propriétés suivantes :

- (1) $N(0) = 0$.
- (2) Le processus est à accroissements indépendants. (le nombre d'évènements se produisant dans deux espaces disjoints sont indépendants)
- (3) Pour $a \leq b$, si on note par $N_{a,b}^d$ le nombre le nombre d'évènements se produisant dans l'espace $V = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_d, b_d]$, alors $N_{a,b}^d$ suit une loi de Poisson de paramètre $\Lambda = \int_V \lambda(t) dt$.

Soient maintenant $N_n(t)$ un processus de Poisson non homogène sur I^d , d'intensité $E(N_n(t)) = nF(t)$ et $Z_n(t) = n^{-1/2}(N_n(t) - nF(t))$ le processus standardisé associé à $N_n(t)$.

Le résultat suivant nous donne un lien entre les variables Z_n et U_n^F

Proposition 2.4. La loi conditionnelle $Z_n(t)|N_n(1) = n$ est la même que celle de $U_n^F(t)$.

On a besoin du lemme suivant :

Lemme 2.5. Soit X et Y deux variables aléatoires indépendantes de loi de Poisson de paramètres respectifs λ et μ . La loi conditionnelle de $X|X + Y = n$ est binomiale de paramètres n et $\frac{\lambda}{\lambda + \mu}$.

Preuve. On sait que $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$, on a alors pour $k \in \{0, 1, \dots, n\}$:

$$\begin{aligned}
 P(X = k | X + Y = n) &= \frac{P(X = k, X + Y = n)}{P(X + Y = n)}, \\
 &= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)}, \\
 &= \frac{\lambda^k \exp(-\lambda)}{k!} \frac{\mu^{n-k} \exp(-\mu)}{(n-k)!} \frac{n!}{(\lambda + \mu)^n \exp(-\lambda - \mu)}, \\
 &= \binom{n}{k} \frac{\lambda^k \mu^{n-k}}{(\lambda + \mu)^n}, \\
 &= \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu} \right)^k \left(1 - \frac{\lambda}{\lambda + \mu} \right)^{n-k}.
 \end{aligned}$$

□

Preuve de la Proposition 2.4. On peut écrire

$$N_n(\mathbf{t})|N_n(\mathbf{1}) = N_n(\mathbf{t})|N_n(\mathbf{t}) + N_n(\mathbf{1}) - N_n(\mathbf{t}),$$

puisque $N_n(\mathbf{t})$ suit une loi de Poisson de paramètre $nF(\mathbf{t})$ indépendante de $N_n(\mathbf{1}) - N_n(\mathbf{t})$ qui est de Poisson de paramètre $n - nF(\mathbf{t})$, en appliquant le Lemme 2.5, on déduit que $N_n(\mathbf{t})|N_n(\mathbf{1}) = n$ suit une binomiale de paramètres n et $F(\mathbf{t})$. Le lemme en découle. \square

Les deux propositions que voici sont fort utiles.

Proposition 2.6. Soit $R \in \mathcal{R}$ vérifiant $F(R) \leq 1/2$, pour tout $\lambda \in \mathbb{R}$ et pour chaque choix du signe on a :

$$P\left(\sup_{S \in \mathcal{R}, \tilde{S} \subset R} \pm U_n^F(S) \geq \lambda\right) \leq 2P\left(\sup_{S \in \mathcal{R}, \tilde{S} \subset R} \pm Z_n(S) \geq \lambda\right). \quad (2.6)$$

Preuve. Soit $A_1 = \{\sup_{S \in \mathcal{R}, \tilde{S} \subset R} \pm U_n^F(S) \geq \lambda\}$ et $A_2 = \{\sup_{S \in \mathcal{R}, \tilde{S} \subset R} \pm Z_n(S) \geq \lambda\}$. Puisque $N_n(R)$ et $N_n(R^c)$ sont indépendants, en utilisant la Proposition 2.4 on a :

$$\begin{aligned} P(A_1) &= P(A_2|N_n(\mathbf{1}) = n), \\ &= \frac{P(A_2, N_n(\mathbf{1}) = n)}{P(N_n(\mathbf{1}) = n)}, \\ &= \sum_{k=0}^n \frac{P(A_2, N_n(R) = k, N_n(R^c) = n - k)}{P(N_n(\mathbf{1}) = n)}, \\ &= \sum_{k=0}^n \frac{P(N_n(R^c) = n - k)}{P(N_n(\mathbf{1}) = n)} P(A_2, N_n(R) = k), \\ &\leq \frac{P(N_n(R^c) = \lfloor nF(R^c) \rfloor)}{P(N_n(\mathbf{1}) = n)} P(A_2). \end{aligned}$$

La formule Stirling $n! \sim (n/e)^n \sqrt{2\pi n}$, quand $n \rightarrow \infty$, nous permet d'écrire

$$\frac{P(N_n(R^c) = \lfloor nF(R^c) \rfloor)}{P(N_n(\mathbf{1}) = n)} \cong \frac{n!}{n^n} e^{n \frac{(nF(R^c))^{\lfloor nF(R^c) \rfloor}}{\lfloor nF(R^c) \rfloor!}} e^{-nF(R^c)} \leq \frac{1}{\sqrt{F(R^c)}}.$$

On déduit le résultat de la condition $F(R) \leq 1/2$. \square

Proposition 2.7. Soit $R \in \mathcal{R}$, pour tout $\lambda \in \mathbb{R}$ et pour chaque choix du signe on a :

$$P\left(\sup_{S \in \mathcal{R}, \tilde{S} \subset R} \pm Z_n(S) \geq \lambda\right) \leq 2^{2d} P\left(\pm Z_n(R) \geq \lambda - 2d\sqrt{2F(R)}\right). \quad (2.7)$$

Preuve. Pour $k \in \mathbb{N}_0$, soit $\mathcal{P}_k \subset \mathcal{R}$, une partition de R en 2^{kd} rectangles de même volume de façon que le i -ième coté d'un élément de \mathcal{P}_k soit égal à 2^{-k} fois le i -ième coté de R . On peut alors voir la collection $\{\pm Z_n(P)\}_{P \in \mathcal{P}_k}$ comme une suite $\{\xi_t\}_{t \leq v}$ à la quelle on peut appliquer la Proposition 2.3 avec

$$\sigma^2 = \sum_{P \in \mathcal{P}_k} \text{Var}(Z_n(P)) = \sum_{P \in \mathcal{P}_k} F(P) = F(R).$$

Puisque

$$\lim_{k \rightarrow \infty} \sup_{S \in \mathcal{A}(R, k)} \pm Z_n(S) = \sup_{S \in \mathcal{R}, S \subset R} \pm Z_n(S), \text{ p. s.} \quad (2.8)$$

où $\mathcal{A}(R, k) = \{S \in \mathcal{R}, S \subset R, S \text{ s'écrit comme une réunion d'éléments de } \mathcal{P}_k\}$, on déduit alors de tout ce qui précède que :

$$\mathbb{P} \left(\sup_{S \in \mathcal{R}, S \subset R} \pm Z_n(S) > \lambda \right) \leq 2^{2d} \mathbb{P} \left(\pm Z_n(R) \geq \lambda - 2d\sqrt{2F(R)} \right) \quad (2.9)$$

et le résultat en découle. \square

2.2. INÉGALITÉS DE BENNETT-WELLNER

Intéressons nous maintenant aux inégalités concernant les variables aléatoires de loi binomiale, particulièrement à celle de Bennett-Wellner. Commençons par définir la fonction suivante.

Soit $\psi : [-1, \infty) \rightarrow [0, \infty)$ définie par :

$$\psi(\lambda) = \begin{cases} 2\lambda^{-2}((1+\lambda)\log(1+\lambda) - \lambda) = 2\lambda^{-2}h(1+\lambda), & \lambda \in (-1, 0) \cup (0, \infty), \\ \psi(0) = 1 \text{ et } \psi(-1) = 2, & \end{cases} \quad (2.10)$$

avec : $h(\lambda) = \lambda(\log \lambda - 1) + 1$.

Proposition 2.8. *Les propriétés principales de la fonction ψ sont les suivantes.*

- (a) ψ est décroissante et continue.
- (b) $\psi(\lambda) \sim 2 \log \lambda / \lambda$ quand $\lambda \rightarrow \infty$.
- (c) $(\cdot)\psi(\cdot)$ est croissante.

Preuve. Pour prouver (a), posons

$$g(\lambda) = \psi(\lambda - 1)/2 = h(\lambda)/(\lambda - 1)^2,$$

on a $g'(\lambda) = (\lambda + 1)f(\lambda)/(\lambda - 1)^3$ où $f(\lambda) = 2(\lambda - 1)/(\lambda + 1) - \log \lambda$.

On a $f'(\lambda) = -(1 - \lambda)^2/(\lambda(1 + \lambda)^2) < 0$ et $f(1) = 0$, donc $f(\lambda) > 0$ pour $\lambda < 1$ (respectivement $f(\lambda) < 0$ pour $\lambda > 1$). Il en découle que $g'(\lambda) < 0$ pour tout $\lambda \neq 1$ et $g'(1) = -1/6$. Et ψ est décroissante pour $\lambda \geq 1$. La partie (b) est évidente. Pour (c), posons

$$g(\lambda) = (\lambda - 1)\psi(\lambda - 1)/2 = h(\lambda)/(\lambda - 1),$$

$g'(\lambda) = ((\lambda - 1) - \log \lambda)/(\lambda - 1)^2$. On a $g'(1) = 1/2$ et $g'(\lambda) > 0$ pour tout $\lambda \neq 1$, on déduit que $(\cdot)\psi(\cdot)$ est croissante. \square

La proposition suivante nous donne l'inégalité de Bennett.

Proposition 2.9. Soient X_1, X_2, \dots, X_n une suite de variables aléatoires indépendantes vérifiant $X_k \leq b$, $E(X_k) = 0$ et $Var(X_k) = \sigma_k^2$, $1 \leq k \leq n$. Alors pour $\sigma^2 = \sum_{k=1}^n \sigma_k^2/n$ et $\lambda \geq 0$ on a :

$$P\left(\sum_{k=1}^n X_k/\sqrt{n} \geq \lambda\right) \leq \exp\left(-\frac{\lambda^2}{2\sigma^2}\psi\left(\frac{\lambda b}{\sigma^2\sqrt{n}}\right)\right).$$

Preuve. En utilisant l'inégalité de Markov on trouve

$$\begin{aligned} P\left(\sum_{k=1}^n X_k \geq \lambda\right) &\leq \inf_{r>0} P\left(\exp\left(r \sum_{k=1}^n X_k\right) \geq \exp(r\lambda)\right), \\ &\leq \inf_{r>0} \exp(-\lambda r) E\left(\exp\left(r \sum_{k=1}^n X_k\right)\right). \end{aligned} \quad (2.11)$$

Posons $g(x) = (e^x - 1 - x)/x^2$ pour $x \neq 0$ et $g(0) = 1/2$. La fonction g est positive croissante et convexe sur \mathbb{R} . On peut alors écrire :

$$\begin{aligned} E(\exp(rX_i)) &= E(1 + rX_i + r^2X_i^2g(rX_i)), \\ &\leq 1 + r^2\sigma_i^2g(rb), \\ &\leq \exp\left(\sigma_i^2 \frac{e^{rb} - 1 - rb}{b^2}\right). \end{aligned} \quad (2.12)$$

En utilisant (2.11), (2.12) et l'indépendance des X_k , $k = 1, \dots, n$ on déduit que :

$$P\left(\sum_{k=1}^n X_k \geq \lambda\right) \leq \inf_{\lambda>0} \exp\left(-r\lambda + n\sigma^2 \frac{e^{rb} - 1 - rb}{b^2}\right). \quad (2.13)$$

En dérivant l'expression du membre de droite de (2.13), on trouve que le minimum est atteint pour $r = b^{-1} \log \left(1 + \frac{\lambda b}{n\sigma^2} \right)$. Cette valeur dans (2.11) donne :

$$\begin{aligned}
 P \left(\sum_{k=1}^n X_k \geq \lambda \right) &\leq \exp \left(-\frac{\lambda}{b} \log \left(1 + \frac{\lambda b}{n\sigma^2} \right) + n\sigma^2 b^{-2} \left[\frac{\lambda b}{n\sigma^2} - \log \left(1 + \frac{\lambda b}{n\sigma^2} \right) \right] \right), \\
 &= \exp \left(-\frac{\lambda}{b} \left[\left(1 + \frac{n\sigma^2}{\lambda b} \right) \log \left(1 + \frac{\lambda b}{n\sigma^2} \right) - 1 \right] \right), \\
 &= \exp \left(-\frac{\lambda}{b} \left[\left(1 + \frac{\lambda b}{n\sigma^2} \right) \log \left(1 + \frac{\lambda b}{n\sigma^2} \right) - \frac{\lambda b}{n\sigma^2} \right] \right) / \frac{\lambda b}{n\sigma^2}, \\
 &= \exp \left(-\frac{\lambda^2}{2n\sigma^2} \psi \left(\frac{\lambda b}{n\sigma^2} \right) \right).
 \end{aligned}$$

□

Posons $X = (\text{Binomiale}(n, p) - np) / \sqrt{n}$ et $q = 1 - p$.

Proposition 2.10. *Pour $0 \leq p \leq 1/2$, pour tout $\lambda > 0$ on a pour chaque choix du signe*

$$P(\pm X \geq \lambda) \leq \exp \left(-\frac{\lambda^2}{2p} \psi \left(\pm \frac{\lambda}{p\sqrt{n}} \right) \right) \quad (2.14)$$

Preuve. Pour le choix du signe positif, on peut écrire $X = \sum_{k=1}^n X_k / \sqrt{n}$ où les $X_k = \text{Bernouilli}(p) - p$ sont indépendantes, $k = 1, \dots, n$. Une application de la Proposition 2.9 avec $b = q$ et $\sigma^2 = pq$, permet de déduire le résultat. Pour choix de signe négatif. Commençons par remarquer que

$$P(-X/\sqrt{n} \leq p) = P(\text{Binomiale}(n, p) \geq 0) = 1.$$

La probabilité dans (2.14) est donc nulle pour $\lambda > \sqrt{np}$, on peut sans nuire à la généralité supposer que $\lambda \leq \sqrt{np}$. Sachant que la fonction génératrice des moments d'une Binomiale(n, p) est $(q + pe^t)^n$ on a :

$$\begin{aligned}
 P(-X \geq \lambda) &\leq P(e^{-rX} \geq e^{r\lambda}), \\
 &\leq e^{-r\lambda} E(e^{-rX}), \\
 &= \exp \left(r(\sqrt{np} - \lambda) + n \log \left(q + pe^{-r/\sqrt{n}} \right) \right). \quad (2.15)
 \end{aligned}$$

En dérivant l'expression dans (2.15), on trouve que le minimum est atteint pour

$$r = \sqrt{n} \left[\log \left(1 + \frac{\lambda}{\sqrt{np}} \right) - \log \left(1 - \frac{\lambda}{\sqrt{nq}} \right) \right],$$

en mettant cette expression dans (2.15) on trouve :

$$P(-X \geq \lambda) \leq \exp(-ng(\lambda/\sqrt{n})), \quad (2.16)$$

où

$$g(\mu) = (p - \mu) \log(1 - \mu/p) + (q + \mu) \log(1 + \mu/q), \quad \mu = \lambda/\sqrt{n}.$$

On a rapidement

$$\begin{aligned} g(\mu) &= \frac{\mu^2}{2p} \psi\left(-\frac{\mu}{p}\right) + \frac{\mu^2}{2q} \psi\left(\frac{\mu}{q}\right), \\ &\geq \frac{\mu^2}{2p} \psi\left(-\frac{\mu}{p}\right). \end{aligned}$$

On déduit alors aisément le résultat de (2.16). \square

Proposition 2.11. *Soit Z une variable aléatoire de loi de Poisson de paramètre τ , alors lorsque $\lambda \geq 0$ on a pour chaque choix du signe*

$$P(\pm(Z - \tau) \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2\tau} \psi\left(\pm\frac{\lambda}{\tau}\right)\right). \quad (2.17)$$

On suppose que $\lambda \leq \tau$, pour le choix du signe négatif.

Preuve. Appliquons la Proposition 2.10 avec $X_n = \text{Binomiale}(n, \tau/n) - \tau$, on a :

$$P(\pm X_n \geq \lambda) = P(X_n/\sqrt{n} \leq \lambda/\sqrt{n}) \leq \exp\left(-\frac{\lambda^2}{2\tau} \psi\left(\pm\frac{\lambda}{\tau}\right)\right).$$

On déduit le résultat en faisant tendre n vers l'infini. \square

2.3. PREUVE DU THÉORÈME 2.1

Nous allons enfin donner quelques résultats qui nous permettront de prouver le Théorème 2.1. La preuve des cas (b), (c) et (d) étant similaire au cas (a), nous allons nous intéresser plus particulièrement aux cas (a) et (e).

On peut maintenant prouver le résultat suivant :

Proposition 2.12. *Soit $R \in \mathcal{R}$ et vérifiant $0 < F(R) \leq 1/2$. Alors pour tout $\varepsilon \in (0, 1)$ on a :*

$$P \left(\sup_{S \in \mathcal{R}} U_n^F(S) \geq \lambda \right) \leq C^+ \exp \left(\frac{-(1-\varepsilon)\lambda^2}{2F(R)} \psi \left(\frac{\lambda}{n^{\frac{1}{2}}F(R)} \right) \right), \lambda \geq 0 \quad (2.18)$$

$$P \left(\sup_{S \in \mathcal{R}} -U_n^F(S) \geq \lambda \right) \leq C^- \exp \left(\frac{-(1-\varepsilon)\lambda^2}{2F(R)} \psi \left(-\frac{(1-\varepsilon)\lambda}{n^{\frac{1}{2}}F(R)} \right) \right), 0 \leq \lambda \leq n^{\frac{1}{2}}F(R) \quad (2.19)$$

$$P \left(\sup_{S \in \mathcal{R}} |U_n^F(S)| \geq \lambda \right) \leq C \exp \left(\frac{-(1-\varepsilon)\lambda^2}{2F(R)} \psi \left(\frac{\lambda}{n^{\frac{1}{2}}F(R)} \right) \right), \lambda \geq 0 \quad (2.20)$$

où C^+ , C^- et $C \in (0, \infty)$ ne dépendent que de ε et d . $S \in \mathcal{R}$.

Preuve. En appliquant successivement les inégalités (2.6), (2.7) et (2.17), on peut écrire pour $\lambda \geq 4d\sqrt{2F(R)}/\varepsilon$

$$\begin{aligned} P \left(\sup_{S \in \mathcal{R}} \pm U_n^F(S) \geq \lambda \right) &\leq 2^{2d+1} P \left(\pm Z_n(R) \geq \lambda - 2d\sqrt{2F(R)} \right), \\ &\leq 2^{2d+1} P \left(\pm Z_n(R) \geq \left(1 - \frac{\varepsilon}{2}\right) \lambda \right), \\ &\leq 2^{2d+1} \exp \left(\frac{-(1 - \frac{\varepsilon}{2})^2 \lambda^2}{2F(R)} \psi \left(\frac{\pm \lambda (1 - \frac{\varepsilon}{2})}{n^{\frac{1}{2}}F(R)} \right) \right), \\ &\leq 2^{2d+1} \exp \left(\frac{-(1 - \varepsilon)\lambda^2}{2F(R)} \psi \left(\frac{\pm \lambda (1 - \frac{\varepsilon}{2})}{n^{\frac{1}{2}}F(R)} \right) \right). \end{aligned} \quad (2.21)$$

On déduit aisément (2.18) ensuite (2.19) et enfin (2.20) de (2.21). \square

Construisons une suite $\{\overline{X}_i\}_{i=1}^\infty$ de variables aléatoires sur I^d en posant :

$$\overline{X}_i = (1 - X_{i1}, \dots, 1 - X_{id}, X_{i1}, \dots, X_{id}), \quad i \in \mathbb{N}.$$

Nous noterons par \overline{F} la fonction de répartition associée à \overline{X}_i et par \overline{F}_n la fonction empirique. Pour $s \in I^d$ posons $s' = \mathbf{1} - s$, notons que si F est la distribution uniforme

sur I^d .

$$\bar{F}(s', t) = \begin{cases} |t - s|, & \text{si } (s', t) \in I^d \times I^d, s \leq t, \\ 0, & \text{sinon,} \end{cases}$$

dans ce cas en notant par $\bar{U}_n^{\bar{F}} = \bar{U}_n$ le processus empirique associé à $\bar{X}_1, \dots, \bar{X}_n$, on a :

$$\bar{U}_n(s', t) = U_n(R(s, t)), \quad s \leq t, \quad s, t \in I^d, \quad R \in \bar{\mathcal{R}}.$$

On voit alors qu'on peut écrire

$$\omega_n(\alpha) = \sup_{\bar{F}(s, t) \leq \alpha} |\bar{U}_n(s, t)| \quad (2.22)$$

On aura besoin des deux propositions suivantes.

Proposition 2.13. *Soit $\varepsilon \in (0, 1/2]$ et $\alpha \in (0, 1/4)$. Alors on a*

$$P(\omega_n(\alpha) \geq \lambda \alpha^{1/2}) \leq \frac{C}{\alpha} \left(\log \frac{1}{\alpha} \right)^{d-1} \exp \left(\frac{-(1-\varepsilon)\lambda^2}{2} \psi \left(\frac{\lambda}{(n\alpha)^{1/2}} \right) \right), \quad \lambda \geq 0 \quad (2.23)$$

avec $C = C(d, \varepsilon) > 0$.

Preuve. Définissons θ par $\theta^{2d+1} = 1 - \varepsilon$ et $l \in \mathbb{N}$ par $\theta^{l+1} < \alpha \leq \theta^l$. On a aisément $l \leq \log(1/\alpha)/\log(1/\theta)$. Pour $k \in \mathbb{N}$ posons

$$A(\theta, k) = \left\{ v \in I^d : v_j = \theta^{k(j)}, \quad k(j) \in \mathbb{N}, \quad \forall j \in \{1, \dots, d\} \text{ et } \sum_{j=1}^d k(j) = k \right\}, \quad (2.24)$$

pour $\beta \in (1, 2)$ et $x \in (0, 1)$.

$$\Psi(\beta, x) = \left\{ (1 - k(\beta - 1)x, \beta x + k(\beta - 1)x) \in I^2 : k \in \mathbb{N}, k \leq \left\lfloor \frac{1 - \beta x}{(\beta - 1)x} \right\rfloor \right\},$$

et

$$B(\beta, x) = \begin{cases} \Psi(\beta, x) \cup \{(\beta x, 1)\} & \text{si } \beta x < 1, \\ \Psi(\beta, x) \cup \{(1, 1)\} & \text{si } \beta x \geq 1. \end{cases} \quad (2.25)$$

On voit aisément que $\#A(\theta, k) = \binom{k+d-1}{d-1}$ et que $\#B(\beta, x) \leq 1/((\beta - 1)x)$. Puisque $\lambda\psi(\lambda)$ est croissante pour $\lambda \geq -1$, une application de (2.20) donne pour $(\tilde{s}, \tilde{t}) \in I^{2d}$ vérifiant $\bar{F}(\tilde{s}, \tilde{t}) \leq \alpha/(1 - \varepsilon)$,

$$P \left(\sup_{(s, t) \leq (\tilde{s}, \tilde{t})} |\bar{U}_n(s, t)| \geq \lambda \alpha^{1/2} \right) \leq C \exp \left(\frac{-(1-\varepsilon)^2 \lambda^2}{2} \psi \left(\frac{\lambda}{(n\alpha)^{1/2}} \right) \right). \quad (2.26)$$

D'après (2.22), (2.24) et (2.25)

$$\begin{aligned}
P(\omega_n(\alpha) \geq \lambda \alpha^{1/2}) &\leq P\left(\sup_{\bar{F}(s,t) \leq \theta^l} |\bar{U}_n(s,t)| \geq \lambda \alpha^{1/2}\right), \\
&\leq P\left(\max_{u \in \Lambda(\theta, l-d)} \sup_{t+s-1 \leq u} |\bar{U}_n(s,t)| \geq \lambda \alpha^{1/2}\right), \\
&\leq \sum_{u \in \Lambda(\theta, l-d)} P\left(\sup_{t+s-1 \leq u} |\bar{U}_n(s,t)| \geq \lambda \alpha^{1/2}\right), \\
&\leq \sum_{u \in \Lambda(\theta, l-d)} P\left(\max_{(\tilde{s}, \tilde{t}) \in \prod_{j=1}^d B\left(\frac{1}{\theta}, u_j\right)} \sup_{(s,t) \leq (\tilde{s}, \tilde{t})} |\bar{U}_n(s,t)| \geq \lambda \alpha^{1/2}\right), \\
&\leq \sum_{u \in \Lambda(\theta, l-d)} \sum_{(\tilde{s}, \tilde{t}) \in \prod_{j=1}^d B\left(\frac{1}{\theta}, u_j\right)} P\left(\sup_{(s,t) \leq (\tilde{s}, \tilde{t})} |\bar{U}_n(s,t)| \geq \lambda \alpha^{1/2}\right).
\end{aligned} \tag{2.27}$$

Puisque pour $(\tilde{s}, \tilde{t}) \in \prod_{j=1}^d B\left(\frac{1}{\theta}, u_j\right)$ on a $\bar{F}(\tilde{s}, \tilde{t}) = \prod_{j=1}^d \left(\frac{u_j}{\theta} \wedge 1\right) \leq \theta^{l-2d} \leq \alpha/(1-\varepsilon)$, on peut appliquer (2.26) à la probabilité apparaissant au membre de droite de l'inégalité (2.27), on obtient sachant que $\binom{l-1}{d-1} \leq l^{d-1}$:

$$\begin{aligned}
P(\omega_n(\alpha) \geq \lambda \alpha^{1/2}) &\leq \binom{l-1}{d-1} \prod_{j=1}^d \frac{\theta}{(1-\theta)u_j} C \exp\left(\frac{-(1-\varepsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}}\right)\right), \\
&= C_0 \left(\log \frac{1}{\theta}\right)^{1-d} \left(\frac{\theta}{1-\theta}\right)^d \frac{1}{\alpha} \left(\log \frac{1}{\alpha}\right)^{d-1} \\
&\quad \times \exp\left(\frac{-(1-\varepsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}}\right)\right).
\end{aligned}$$

On déduit alors le résultat en posant $C = C_0 \left(\log \frac{1}{\theta}\right)^{1-d} \left(\frac{\theta}{1-\theta}\right)^d$ et en remplaçant de nouveau $(1-\varepsilon)$ par $(1-\varepsilon)^2$. \square

Proposition 2.14. Soit $\varepsilon \in (0, 1)$ et $\alpha \in (0, 1]$. Posons $n_k = \left\lfloor (1 + \varepsilon/2)^k \right\rfloor$, $k \in \mathbb{N}$. Pour $k \in \mathbb{N}$ et $\lambda > 2(\alpha/\varepsilon)^{\frac{1}{2}}$, on a

$$P\left(\sup_{n_k < n \leq n_{k+1}} \omega_n(\alpha) \geq \lambda\right) \leq 2P\left(\omega_{n_{k+1}}(\alpha) \geq \lambda(1-\varepsilon)\right). \tag{2.28}$$

On a besoin du lemme suivant

Lemme 2.15. Soit I un ensemble d'indice muni d'un bon ordre (i.e totalement ordonné et chaque sous ensemble possède un plus petit élément) et pour $m \in I$, soit A_m et B_m des évènements d'un espace probabilisé. Si pour tout $m \in I$, $A_m \cap \left(\bigcap_{k < m} A_k^c\right)$ et B_m sont indépendants, alors

$$\inf_{m \in I} P(B_m) P\left(\bigcup_{m \in I} A_m\right) \leq P\left(\bigcup_{m \in I} (A_m \cap B_m)\right). \quad (2.29)$$

Preuve. Soit $<$ un bon ordre sur I , on peut écrire (avec $C_m = A_m \cap B_m$)

$$\begin{aligned} P\left\{\bigcup_{m \in I} C_m\right\} &= P\left\{\bigcup_{m \in I} \left(\left(\bigcap_{k < m} C_k^c\right) \cap C_m\right)\right\}, \\ &= \sum_{m \in I} P\left\{\left(\bigcap_{k < m} C_k^c\right) \cap C_m\right\}, \\ &\geq \sum_{m \in I} P\left\{A_m \cap \left(\bigcap_{k < m} A_k^c\right) \cap B_m\right\}, \\ &= \sum_{m \in I} P\left\{A_m \cap \left(\bigcap_{k < m} A_k^c\right)\right\} P\{B_m\}, \\ &\geq \inf_{m \in I} P(B_m) \sum_{m \in I} P\left\{A_m \cap \left(\bigcap_{k < m} A_k^c\right)\right\}, \\ &= \inf_{m \in I} P(B_m) P\left(\bigcup_{m \in I} A_m\right). \end{aligned}$$

□

Preuve de Proposition 2.14. Soit $\{r_i, i \in \mathbb{N}\}$ une énumération de

$$\{(s, t) \in I^d \times I^d : \bar{F}(s, t) \leq \alpha\} \cap \mathbb{Q}^{2d},$$

et posons

$$S_n(s, t) = n^{1/2} \bar{U}_n(s, t), \quad (s, t) \in I^d \times I^d.$$

On peut alors écrire

$$P\left(\sup_{n_k < n \leq n_{k+1}} \omega_n(\alpha) > \lambda\right) \leq P\left(\sup_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} |S_n(r_i)| > \lambda(n_k + 1)^{1/2}\right) \quad (2.30)$$

Pour $n_k < n \leq n_{k+1}$ et $i \in \mathbb{N}$, posons

$$\begin{aligned} A_{ni} &= \left\{|S_n(r_i)| > \lambda(n_k + 1)^{1/2}\right\}, \\ B_{ni} &= \left\{|S_{n_{k+1}}(r_i) - S_n(r_i)| > \lambda(n_k + 1)^{1/2} \times \frac{\varepsilon}{2}\right\}. \end{aligned}$$

En appliquant (2.29) sur l'ensemble $\{(n, i) : n_k < n \leq n_{k+1}, i \in \mathbb{N}\}$ muni de l'ordre lexicographique, on a que la probabilité à droite en (2.30) est plus petite ou égale à

$$\left(\inf_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} P(B_{ni}) \right)^{-1} P\left(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| > (1 - \frac{\varepsilon}{2}) \lambda (n_k + 1)^{1/2} \right).$$

Il est évident de voir que

$$P\left(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| > \lambda (1 - \varepsilon/2) (n_k + 1)^{1/2} \right) \leq P\left(\sup_{\bar{F}(s,t) \leq \alpha} |\bar{U}_{n_{k+1}}(s, t)| \geq \lambda (1 - \varepsilon) \right).$$

Montrons enfin que $P(B_{ni}) \geq 1/2$, pour $n_k < n \leq n_{k+1}$ et $i \in \mathbb{N}$. Étant donné que $\lambda > 2(\alpha/\varepsilon)^{1/2}$, en appliquant l'inégalité de Markov on a :

$$P(B_{ni}^c) \leq \frac{4\bar{F}(r_i)(1 - \bar{F}(r_i))(n_{k+1} - (n_k + 1))}{\lambda^2(n_k + 1)\varepsilon^2} \leq \frac{2\varepsilon\alpha}{\lambda^2\varepsilon^2} \leq \frac{1}{2}.$$

□

Nous sommes maintenant en mesure de prouver le Théorème 2.1.

Preuve du Théorème 2.1. Commençons par prouver (a). Soit $\varepsilon \in (0, 1/2)$, posons

$$A_k = \left\{ \omega_{n_{k+1}}(a_{n_k}) \geq (2(1 + 2\varepsilon)(1 - \varepsilon)^{-1}(1 + c)a_{n_{k+1}} \log \log n_k)^{1/2} \right\}.$$

Pour montrer que

$$\limsup_{n \rightarrow \infty} \frac{\omega_n(a_n)}{(a_n \log \log n)^{1/2}} \leq (2(1 + c))^{1/2} \text{ p.s.}, \quad (2.31)$$

d'après la Proposition 2.14 et le lemme de Borel-Cantelli, il suffit de montrer que pour un $k_0 \in \mathbb{N}_0$, $\sum_{k \geq k_0} P(A_k)$ converge. Une application de la Proposition 2.13 avec $\alpha = a_{n_k} \leq 1/4$ (pour $k \geq k_0$, $k_0 \in \mathbb{N}_0$) donne

$$\begin{aligned} P(A_k) &\leq C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}} \right)^{d-1} \exp \left(- (1 + 2\varepsilon)(1 + c) \log \log n_k \frac{a_{n_{k+1}}}{a_{n_k}} \right. \\ &\quad \left. \times \psi \left(\frac{(2(1 + 2\varepsilon)(1 - \varepsilon)^{-1}(1 + c)a_{n_{k+1}} \log \log n_k)^{1/2}}{(n_{k+1} a_{n_k})^{1/2}} \right) \right) \end{aligned} \quad (2.32)$$

Les hypothèses en (a) nous montrent qu'il existe $k_1 \in \mathbb{N}$ tel que $k \geq k_1$,

$$a_{n_{k+1}} \geq \frac{1}{(\log n_k)^{\varepsilon+c}}, \quad (2.33)$$

en prenant $k \geq \max(k_0, k_1)$, étant donné (2.33), le terme à l'intérieur de ψ dans (2.32) tend vers 0 pour k tendant vers l'infini. Donc

$$\begin{aligned} P(A_k) &\leq C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}} \right)^{d-1} \exp \left(- (1 + 2\varepsilon)(1 + c) \log \log n_k \right), \\ &= C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}} \right)^{d-1} \left(\frac{1}{\log n_k} \right)^{(1+\varepsilon)(1+c)}, \end{aligned}$$

en utilisant à nouveau (2.33) on a rapidement

$$P(A_k) \leq C(c + \varepsilon)^{d-1} \frac{(\log \log n_k)^{d-1}}{(\log n_k)^{1+c\varepsilon}}.$$

On déduit que

$$P(A_k) \leq C(c + \varepsilon)^{d-1} \frac{\left(\log k + \log \log(1 + \varepsilon/2) \right)^{d-1}}{\log(1 + \varepsilon/2)^{1+c\varepsilon} k^{1+c\varepsilon}}.$$

Le résultat (2.31) est obtenu du critère de convergence des séries de Bertrand.

Nous renvoyons à Einmahl (1987) pour la preuve du cas (e).

□

Chapitre 3

ESTIMATEURS DE BERNSTEIN

Dans ce chapitre nous revenons au cas univarié $d = 1$. On désignera donc par $X_1, X_2, \dots, X_n, \dots$ une suite de variables aléatoires définies sur un espace probabilisé (Ω, \mathcal{F}, P) à support contenu dans $I = [0, 1]$. Soit F la fonction de répartition commune des éléments de la suite $\{X_i\}_{i=1}^\infty$. On suppose que F est continue et on note par f la densité commune.

L'estimateur de Bernstein étant lié à la fonction de répartition empirique F_n , nous commençons par donner les principaux résultats concernant la fonction de répartition empirique.

3.1. LA FONCTION DE RÉPARTITION EMPIRIQUE

Rappelons la définition de la fonction de répartition empirique (2.1) :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{[X_i \leq x]},$$

pour $x \in I$. Il est évident que $nF_n(x)$ suit une loi binomiale de paramètres n et $F(x)$. On en déduit que $F_n(x)$ est un estimateur sans biais de $F(x)$ et convergent, en effet $\text{Var}(F_n) = F(x)(1 - F(x))/n$.

Le résultat suivant est un cas particulier du théorème de Glivenko-Cantelli.

Théorème 3.1. *On a le résultat suivant,*

$$\lim_{n \rightarrow \infty} \|F_n - F\| = 0 \quad p.s. \quad (3.1)$$

Pour prouver le Théorème 3.1, on a besoin du lemme suivant.

Lemme 3.2. Soit X une variable aléatoire réelle de fonction de répartition F définie sur I . Pour tout $p \in (0, 1]$, on appelle quantile ou fractile d'ordre p de X le nombre

$$x_p = \inf \{x \in I, F(x) \geq p\}. \quad (3.2)$$

L'application $p \in (0, 1] \mapsto x_p$ est appelée inverse généralisée de F et est notée F^{-1} . De plus si ξ est une variable aléatoire de loi uniforme sur I ,

(i) on a l'équivalence $F(x) \geq p \Leftrightarrow x \geq F^{-1}(p)$.

(ii) En outre si F est continue, alors

$$\text{pour tout } p \in (0, 1), F(F^{-1}(p)) = p. \quad (3.3)$$

La variable $F^{-1}(\xi)$ a même loi que X . Et donc $F(X)$ suit une loi uniforme sur I .

Preuve. D'après (3.2) si $F(x) \geq p$, on a $x \geq F^{-1}(p)$. Réciproquement, pour tout $n \in \mathbb{N}_0$, il existe $x \in I$, $F(x) \geq p$ telle que $F^{-1}(p) + 1/n > x$. De la croissance de F , on déduit que $F(F^{-1}(p) + 1/n) \geq F(x) \geq p$. La fonction F étant continue à droite, en faisant tendre n vers l'infini on a :

$$\text{pour tout } p \in (0, 1), F(F^{-1}(p)) \geq p. \quad (3.4)$$

En utilisant à nouveau la croissance de F , on déduit alors que si $x \geq F^{-1}(p)$ alors $F(x) \geq F(F^{-1}(p)) \geq p$, ce qui prouve (i). Maintenant pour $x < F^{-1}(p)$ on a $F(x) < p$ d'après (i), si F est continue en $F^{-1}(p)$ alors en faisant tendre x vers $F^{-1}(p)$ on a $F(F^{-1}(p)) \leq p$, $p \in (0, 1)$. On déduit alors (3.3) de (3.4). Enfin, pour $x \in I$,

$$P(F^{-1}(\xi) \leq x) = P(\xi \leq F(x)) = F(x),$$

on déduit que les variables aléatoires $F^{-1}(\xi)$ et X suivent la même loi. Et par conséquent $F(X)$ a même loi que $F(F^{-1}(\xi))$ qui est égal à ξ presque partout d'après (3.3). Ce qui termine la preuve du (ii). \square

Preuve du Théorème 3.1. Notons par $\Gamma_n(t)$ la fonction de répartition empirique de n variables aléatoires $\xi_1, \xi_2, \dots, \xi_n$ i.i.d. de loi uniforme sur I . Enfin posons $\Gamma(t) = t$ la densité de ξ_i . Vu (i) et (ii) du Lemme 3.2, on peut alors écrire pour $t \in I$:

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq t\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{F(X_i) \leq F(t)\}} = \Gamma_n(F(t)).$$

On déduit alors rapidement que

$$\|F_n - F\| = \sup_{t \in I} |\Gamma_n(F(t)) - F(t)| = \sup_{x \in F(I)} |\Gamma_n(x) - x| \leq \|\Gamma_n - \Gamma\|.$$

Il suffit donc de montrer le théorème dans le cas où X_i suit une loi uniforme. D'après la loi des grands nombres, on sait que pour x fixé, $\Gamma_n(x)$ converge vers x p.s. Il existe donc une ensemble A_x avec $P(A_x) = 0$ tel que pour tout $\omega \in \Omega \setminus A_x$,

$$\Gamma_n(x)(\omega) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\xi_i(\omega) \leq x\}}$$

converge vers x . On déduit alors qu'il existe A , avec $P(A) = 0$ tel que pour tout $x \in I \cap \mathbb{Q}$ et $\omega \in \Omega \setminus A$, $\Gamma_n(x)(\omega)$ converge vers x . Mais Γ_n étant croissante, le résultat précédent est vrai pour tout $x \in I$. En effet soit $x \in I$ et $\omega \in \Omega \setminus A$, il existe deux suites (x_k) et (y_k) d'éléments de $I \cap \mathbb{Q}$ vérifiant $x_n \nearrow x$ et $y_n \searrow x$ telles que $\Gamma_n(x_k)(\omega) \leq \Gamma_n(x)(\omega) \leq \Gamma_n(y_k)(\omega)$, $k \in \mathbb{N}$ et $n \geq 1$ fixés. En faisant tendre n vers l'infini, on $x_k \leq \lim_{n \rightarrow \infty} \Gamma_n(x)(\omega) \leq y_k$. En faisant enfin tendre k vers l'infini on déduit que $\lim_{n \rightarrow \infty} \Gamma_n(x)(\omega) = x$. En appliquant le premier théorème de Dini (Théorème A.3), on déduit que la convergence est uniforme, et le résultat en découle. \square

Posons pour la suite,

$$a_n = \left(\frac{2n}{\log \log n} \right)^{1/2}.$$

Définition 3.1. Un estimateur \hat{F}_n de F est dit vérifier la propriété de Chung-Smirnov si

$$\limsup_{n \rightarrow \infty} a_n \|\hat{F}_n - F\| \leq 1 \quad \text{p. s.} \quad (3.5)$$

Si

$$\limsup_{n \rightarrow \infty} a_n \|\hat{F}_n - F\| = 1 \quad \text{p. s.} \quad (3.6)$$

on dit alors que \hat{F}_n vérifie la loi du logarithme itéré.

L'inégalité de Chung-Smirnov nous donne une idée précise de l'ordre de convergence de \hat{F}_n vers la véritable fonction F .

Le théorème suivant nous montre que F_n vérifie la loi du logarithme itéré.

Théorème 3.3. On a le résultat suivant,

$$\limsup_{n \rightarrow \infty} a_n \|F_n - F\| = 1 \quad \text{p. s.} \quad (3.7)$$

Preuve. La preuve de ce théorème est assez délicate. Nous renvoyons au Théorème B, section II.1.4 (P. 62) de Serfling (1980). \square

3.2. ESTIMATEURS DE BERNSTEIN

3.2.1. Définitions et Propriétés

En s'inspirant que la convergence en (1.4), Babu et al. (2002) proposent l'estimateur suivant :

$$\tilde{F}_{n,m}(x) = \sum_{k=0}^m F_n(k/m) b_k(m, x), \quad x \in I.$$

La fonction $\tilde{F}_{n,m}$ est un polynôme en x dont les coefficients dépendent des variables aléatoires X_1, X_2, \dots, X_n , elle est donc continue sur I ainsi que toutes ses dérivées. Clairement $0 \leq F_{n,m}(x) \leq 1$ pour $x \in I$. On peut écrire

$$\tilde{F}_{n,m}(x) = \sum_{k=0}^m f_n(k/m) \mathcal{B}_k(m, x), \quad (3.8)$$

où

$$f_n(0) = 0, \quad f_n(k/m) = F_n(k/m) - F_n((k-1)/m), \quad k = 1, 2, \dots, m, \quad (3.9)$$

et

$$\mathcal{B}_k(m, x) = \sum_{j=k}^m b_j(m, x). \quad (3.10)$$

En effet en utilisant (3.10) et vu que $F_n(0) = 0$ pour tout $n \in \mathbb{N}_0$, on a :

$$\begin{aligned} \sum_{k=0}^m f_n(k/m) \mathcal{B}_k(m, x) &= \sum_{k=1}^m F_n(k/m) \mathcal{B}_k(m, x) - \sum_{k=1}^m F_n((k-1)/m) \mathcal{B}_k(m, x), \\ &= F_n(m/m) \mathcal{B}_m(m, x) + \sum_{k=1}^{m-1} F_n(k/m) (\mathcal{B}_k(m, x) - \mathcal{B}_{k+1}(m, x)), \\ &= F_n(m/m) b_m(m, x) + \sum_{k=1}^{m-1} F_n(k/m) b_k(m, x), \\ &= \tilde{F}_{n,m}(x). \end{aligned}$$

En dérivant (3.10) on montre aisément que :

$$\frac{d}{dx} \mathcal{B}_k(m, x) = m \binom{m-1}{k-1} x^{k-1} (1-x)^{m-k} = m b_{k-1}(m-1, x) \geq 0. \quad (3.11)$$

Puisque $f_n(k/m) \geq 0$, pour tout $k = 0, 1, \dots, m$, on déduit que $\tilde{F}_{n,m}$ est une fonction croissante. Et par conséquent $\tilde{F}_{n,m}$ est une fonction de répartition, ce qui justifie son utilisation comme estimateur de la fonction de répartition F . Pour $x \in I$, vu les propriétés de F_n on a rapidement $E[\tilde{F}_{n,m}(x)] = B_m(F)(x)$, pour tout $n \in \mathbb{N}_0$. On déduit donc que $\tilde{F}_{n,m}(x)$ est un estimateur biaisé de $F(x)$, mais asymptotiquement non biaisé.

Dans toute la suite du chapitre on suppose que m est une fonction de n telle que $m = m_n \rightarrow \infty$ lorsque $n \rightarrow \infty$. Nous utiliserons la notation $o_x(\cdot)$ (respectivement $O_x(\cdot)$) pour dire que la limite (respectivement la borne) dépend du point x .

Le théorème suivant nous donne un développement asymptotique des deux premiers moments de l'estimateur $\tilde{F}_{n,m}$.

Théorème 3.4. *Si F est deux fois continûment dérivable à dérivées bornées sur I , alors on a :*

$$E[\tilde{F}_{n,m}(x)] = F(x) + m^{-1}b_1(x) + o(m^{-1}), \quad (3.12)$$

$$\text{Var}[\tilde{F}_{n,m}(x)] = n^{-1}\sigma^2(x) - m^{-1/2}n^{-1}V(x) + o_x(m^{-1/2}n^{-1}), \quad (3.13)$$

où

$$\sigma^2(x) = F(x)(1 - F(x)), \quad b_1(x) = x(1 - x)f'(x)/2 \quad \text{et} \quad V(x) = f(x)(2x(1 - x)/\pi)^{1/2}.$$

Commençons par donner quelques définitions. Posons pour $x \in I$

$$S_m(x) = \sum_{k=0}^m b_k^2(m, x),$$

et pour $j=0, 1$ et 2 ,

$$R_j(m, x) = m^{-j} \sum_{0 \leq k < l \leq m} (k - mx)^j b_k(m, x) b_l(m, x).$$

On peut prouver le lemme suivant

Lemme 3.5. *Soient $\gamma_1(x) = (4\pi x(1 - x))^{-1/2}$ et $\gamma_2(x) = (x(1 - x)/(2\pi))^{1/2}$, on a les résultats suivants :*

- (i) $S_m(x) = m^{-1/2}[\gamma_1(x) + o_x(1)]$, pour $x \in (0, 1)$,
- (ii) $R_1(m, x) = m^{-1/2}[-\gamma_2(x) + o_x(1)]$, pour $x \in (0, 1)$,
- (iii) $0 \leq R_2(m, x) \leq (4m)^{-1}$, pour $x \in (0, 1)$.

Preuve. On voit rapidement que $R_2(m, x)$ est une fonction positive et d'après (1.1a) et (1.1c) on a :

$$R_2(m, x) \leq m^{-2} \sum_{k=0}^m \sum_{l=0}^m (k - mx)^2 b_k(m, x) b_l(m, x) = m^{-2} T_2(m, x) = m^{-1} x(1 - x).$$

Ce qui prouve (iii).

Soient $U_i, V_j, i, j = 1, 2, \dots, m$ des variables aléatoires i.i.d. de loi de Bernoulli de paramètre x . Posons $R_i = (U_i - V_i)/\sqrt{2x(1-x)}$, $U = \sum_{k=0}^m U_k$ et $V = \sum_{k=0}^m V_k$. On a $E(R_i) = 0$, $\text{Var}(R_i) = 1$ et

$$P\left(\sum_{k=0}^m R_k = 0\right) = P(U = V) = \sum_{j=0}^m P(U = j, V = j).$$

Puisque les variables aléatoires U et V suivent des lois binomiales de paramètres m et x indépendantes, on déduit que

$$P\left(\sum_{k=0}^m R_k = 0\right) = \sum_{k=0}^m b_k^2(m, x) = S_m(x).$$

La variable aléatoire R_i est de type "treillis" avec $b = 0$ et $h = 1/\sqrt{x(1-x)}$. En effet R_i prend les valeurs : $0, \pm h$. Une application directe du Théorème A.1 avec $k = 0$ nous permet alors de déduire (i).

D'après le théorème 1 de Cressie (1978), on peut écrire

$$\sum_{l=k}^m b_l(m, x) = 1 - \phi(\delta_k - G_x(\delta_{k-1/2})) + O_x(m^{-1})$$

où

$$\delta_k = (k - mx)(mx(1-x))^{-1/2}, \quad G_x(t) = \left(\frac{1}{2} + \frac{(1-2x)(t^2-1)}{6}\right)(mx(1-x))^{-1/2},$$

et ϕ étant la fonction de répartition d'une variable aléatoire de loi normale centrée réduite. La formule de Taylor donne autour de $t = 0$,

$$\phi(t) = \frac{1}{2} + \frac{t}{\sqrt{2\pi}} + o(|t|).$$

On déduit alors que

$$\sum_{l=k+1}^m b_l(m, x) = \frac{1}{2} - \frac{\delta_{k+1} - G_x(\delta_{k+1/2})}{\sqrt{2\pi}} + o_x(|\delta_{k+1} - G_x(\delta_{k+1/2})|) + O_x(m^{-1}). \quad (3.14)$$

Un calcul simple nous montre que

$$\begin{aligned}
 \delta_{k+1} - G_x(\delta_{k+1/2}) &= \frac{1}{3}(2-x)(mx(1-x))^{-1/2} + \left[1 - \frac{1}{6}(1-2x)(mx(1-x))^{-1}\right] \delta_k \\
 &\quad - \frac{1}{6}(1-2x)(mx(1-x))^{-1/2} \delta_k^2 - \frac{1}{24}(1-2x)(mx(1-x))^{-3/2}, \\
 &= \frac{1}{3}(2-x)(mx(1-x))^{-1/2} + \left[1 - \frac{1}{6}(1-2x)(mx(1-x))^{-1}\right] \delta_k \\
 &\quad - \frac{1}{6}(1-2x)(mx(1-x))^{-1/2} \delta_k^2 + O_x(m^{-3/2}). \tag{3.15}
 \end{aligned}$$

En utilisant (3.14) et (3.15), on peut alors écrire en négligeant les termes d'ordre plus grand ou égal à $O_x(m^{-5/2})$,

$$\begin{aligned}
 R_1(m, x) &= m^{-1} \sum_{k=0}^m (k - mx) b_k(m, x) \left[\sum_{l=k+1}^m b_l(m, x) \right], \\
 &= m^{-1} \left[\frac{1}{2} - \frac{1}{3}(2-x)(2\pi mx(1-x))^{-1/2} \right] T_1(m, x), \\
 &\quad - m^{-1} (2\pi mx(1-x))^{-1/2} T_2(m, x) \\
 &\quad + o_x(m^{-3/2} H_1(m, x)) + o_x(m^{-3/2} H_2(m, x)) + o_x(m^{-5/2} H_3(m, x)),
 \end{aligned}$$

où

$$H_s(m, x) = \sum_{k=0}^m |k - mx|^s b_k(m, x).$$

D'après (1.1b) et (1.1c), on a alors

$$R_1(m, x) = -x(1-x)(2\pi mx(1-x))^{-1/2} + o_x(m^{-1/2}) + o_x(m^{-3/2} H_1(m, x)) + o_x(m^{-5/2} H_3(m, x)).$$

L'équation (1.8) nous montre que $H_1(m, x) \leq \frac{m^{1/2}}{2}$, on a rapidement

$$m^{-3/2} H_1(m, x) \leq m^{-3/2} \frac{m^{1/2}}{2} = (2m)^{-1/2} = O(m^{-1/2}).$$

L'inégalité de Cauchy-Schwarz, donne

$$m^{-5/2} H_3(m, x) \leq m^{-5/2} (T_2(m, x) T_4(m, x))^{1/2} = C m^{-5/2} m^{3/2} = O(m^{-1}).$$

Et on déduit le point (ii). □

Preuve du Théorème 3.4. On obtient aisément (3.12) de (1.11) en prenant la fonction $u = F$. Pour $x \in I$, posons :

$$\Delta_i(x) = \mathbb{1}_{\{X_i \leq x\}} - F(x),$$

on voit qu'on peut alors écrire

$$\tilde{F}_{n,m}(x) - B_m(F)(x) = \sum_{k=0}^m (F_n(k/m) - F(k/m)) b_k(m, x) = \frac{1}{n} \sum_{i=1}^n Y_{i,m},$$

avec

$$Y_{i,m} = \sum_{k=0}^m \Delta_i(k/m) b_k(m, x). \quad (3.16)$$

Pour m fixé, les variables aléatoires $Y_{i,m}$, $i = 1, 2, \dots, n$ sont indépendantes et de moyenne nulle. On déduit que $\text{Var}(\tilde{F}_{n,m}(x)) = \frac{1}{n} E(Y_{1,m}^2)$. Pour $x, y \in I$, on vérifie aisément que

$$E(\Delta_1(x)\Delta_1(y)) = \min(F(x), F(y)) - F(x)F(y).$$

On peut alors écrire

$$\begin{aligned} E(Y_{1,m}^2) &= \sum_{k=0}^m \sum_{l=0}^m E(\Delta_1(k/m)\Delta_1(l/m)) b_k(m, x) b_l(m, x), \\ &= \sum_{k=0}^m F(k/m) b_k^2(m, x) + 2 \sum_{0 \leq k < l \leq m} F(k/m) b_k(m, x) b_l(m, x) \\ &\quad - B_m^2(F)(x). \end{aligned} \quad (3.17)$$

L'idée est maintenant de simplifier cette dernière expression. Pour le premier terme de (3.17), prenons $0 \leq k \leq m$, la formule de Taylor,

$$F(k/m) = F(x) + O(|k/m - x|),$$

nous permet d'avoir

$$\sum_{k=0}^m F(k/m) b_k^2(m, x) = F(x) S_m(x) + O(I_m(x)), \quad (3.18)$$

avec

$$I_m(x) = \sum_{k=0}^m |k/m - x| b_k^2(m, x).$$

Pour le second terme de (3.17), la formule de Taylor nous permet d'écrire

$$F(k/m) = F(x) + (k/m - x)f(x) + O((k/m - x)^2), \quad (3.19)$$

il est aisé de voir que

$$1 = \sum_{k=0}^m \sum_{l=0}^m b_k(m, x) b_l(m, x) = 2R_0(m, x) + S_m(x).$$

On déduit donc que

$$R_0(m, x) = \frac{1}{2}(1 - S_m(x)). \quad (3.20)$$

En utilisant (3.19), (3.20) et (iii) du Lemme 3.10, on peut écrire

$$\begin{aligned} \sum_{0 \leq k < l \leq m} F(k/m) b_k(m, x) b_l(m, x) &= F(x) R_0(m, x) + f(x) R_1(m, x) + O(R_2(m, x)), \\ &= \frac{1}{2} F(x) (1 - S_m(x)) + f(x) R_1(m, x) + O(m^{-1}). \end{aligned} \quad (3.21)$$

En utilisant (3.12), (3.18) et (3.21) on peut réécrire (3.17) sous la forme

$$E(Y_{1,m}^2) = \sigma^2(x) + 2f(x) R_1(m, x) + O(I_m(x)) + O(m^{-1}). \quad (3.22)$$

Puisque $0 \leq b_k(m, x) \leq 1$ pour $0 \leq k \leq m$, l'inégalité de Cauchy-Schwarz nous donne

$$\begin{aligned} I_m(x) &= \sum_{k=0}^m |k/m - x| b_k^{1/2}(m, x) b_k^{3/2}(m, x), \\ &\leq \left(\sum_{k=0}^m (k/m - x)^2 b_k(m, x) \right)^{1/2} \left(\sum_{k=0}^m b_k^3(m, x) \right)^{1/2}, \\ &\leq \left(\frac{T_2(m, x) S_m(x)}{m^2} \right)^{1/2}. \end{aligned}$$

D'après (1.1c) et (i) du Lemme 3.10 on déduit que $I_m(x) = O(m^{-3/4})$. Finalement en utilisant (ii) du Lemme 3.10 on peut réécrire (3.22) sous la forme

$$E(Y_{1,m}^2) = \sigma^2(x) - m^{-1/2} V(x) + o_x(m^{-1/2}). \quad (3.23)$$

On déduit aisément (3.13) de (3.23). \square

Définition 3.2. Soit \hat{F}_n un estimateur de F . On définit l'erreur quadratique moyenne (EQM) de l'estimateur \hat{F}_n au point $x \in I$, par :

$$EQM(\hat{F}_n(x)) = E \left[(\hat{F}_n(x) - F(x))^2 \right], \quad (3.24)$$

Et l'erreur quadratique moyenne intégrée (EQMI) par :

$$EQMI(\hat{F}_n) = E \left[\int_I (\hat{F}_n(x) - F(x))^2 dx \right] = \int_I EQM(\hat{F}_n(x)) dx. \quad (3.25)$$

L'EQM de l'estimateur $\tilde{F}_{n,m}(x)$ est donc

$$\text{EMQ}(\tilde{F}_{n,m}(x)) = n^{-1}\sigma^2(x) - n^{-1}m^{-1/2}V(x) + m^{-2}b_1^2(x) + o_x(n^{-1}m^{-1/2}) + o(m^{-2}). \quad (3.26)$$

Il est aisé de voir que

$$\text{EMQ}(F_n(x)) = \text{Var}(F_n(x)) = n^{-1}\sigma^2(x),$$

les estimateurs $\tilde{F}_{n,m}(x)$ et $F_n(x)$ sont donc équivalents en moyenne quadratique au premier ordre. Le second terme dans (3.26) étant négatif, il pourrait être possible de choisir m de tel sorte que $\tilde{F}_{n,m}(x)$ soit meilleur que $F_n(x)$ en terme d'EQM. Notons qu'il n'est pas possible d'obtenir l'EQMI de l'estimateur $\tilde{F}_{n,m}$ en intégrant l'équation (3.26), ceci est dû à la présence de borne non uniforme en x .

Le corollaire suivant nous donne un choix asymptotique de m .

Corollaire 3.6. *Si $f(x) \neq 0$ et $f'(x) \neq 0$ pour $x \in (0, 1)$, alors sous les hypothèses du Théorème 3.4, le choix optimal de m minimisant asymptotiquement l'EQM est*

$$m_{\text{opt}} = n^{2/3} \left[\frac{4b_1^2(x)}{V(x)} \right]^{2/3},$$

et dans ce cas

$$\text{EMQ}(\tilde{F}_{n,m_{\text{opt}}}(x)) = n^{-1}\sigma^2(x) - n^{-4/3} \frac{3}{4} \left[\frac{V^4(x)}{4b_1^2(x)} \right]^{1/3} + o_x(n^{-4/3}).$$

Preuve. Le résultat est simple à obtenir. L'optimum est trouvé en négligeant les deux termes d'erreur dans (3.26). \square

Comme la fonction de distribution empirique F_n , l'estimateur de Bernstein possède les propriétés suivantes :

Théorème 3.7. *On a le résultat suivant,*

$$\lim_{n, m \rightarrow \infty} \|\tilde{F}_{n,m} - F\| = 0 \quad p.s. \quad (3.27)$$

Preuve. On peut écrire :

$$\|\tilde{F}_{n,m} - F\| \leq \|\tilde{F}_{n,m} - B_m(F)\| + \|B_m(F) - F\|. \quad (3.28)$$

En utilisant (1.1a), on a pour $x \in I$:

$$\begin{aligned}
 |\tilde{F}_{n,m}(x) - B_m(F)(x)| &= \left| \sum_{k=0}^m (F_n(k/m) - F(k/m)) b_k(m, x) \right|, \\
 &\leq \sum_{k=0}^m |F_n(k/m) - F(k/m)| b_k(m, x), \\
 &\leq \max_{0 \leq k \leq m} |F_n(k/m) - F(k/m)| \sum_{k=0}^m b_k(m, x), \\
 &\leq \|F_n - F\|.
 \end{aligned}$$

On en déduit que :

$$\|\tilde{F}_{n,m} - B_m(F)\| \leq \|F_n - F\|, \quad (3.29)$$

et le résultat découle de (1.4), (3.1) et (3.28). \square

Dans la suite posons :

$$b_{n,m} = (n^{-1} \log n)^{1/2} (m^{-1} \log m)^{1/4}, \quad c_n = (n^{-1} \log n)^{1/2}.$$

Le théorème suivant nous permet d'apprécier l'écart entre F_n et $\tilde{F}_{n,m}$.

Théorème 3.8. *Si $f \in \text{Lip}(1)$. Alors pour $n^{2/3} \leq m \leq (n \log n)^2$, on a pour $n \rightarrow \infty$,*

$$\|\tilde{F}_{n,m} - F_n\| = O(b_{n,m}) \quad p.s. \quad (3.30)$$

On a besoin des lemmes suivants.

Lemme 3.9. *Soit Z_1, Z_2, \dots, Z_m une suite de variables aléatoires de moyenne nulle et telle que $|Z_i| \leq b$ pour un $b > 0$, $i = 1, 2, \dots, m$. Si $V \geq \sum_{i=1}^m E(Z_i^2)$, alors pour tout $0 < s < 1$ et $0 \leq a \leq V/(sb)$, on a :*

$$P\left(\sum_{i=1}^m Z_i > a\right) \leq \exp(-a^2 s(1-s)/V). \quad (3.31)$$

Preuve. On montre aisément que $\exp x \leq 1 + x + x^2$ pour $|x| \leq 1$. Pour $0 < \mu < 1/b$, on peut écrire en utilisant l'inégalité de Markov :

$$\begin{aligned}
P\left(\sum_{i=1}^m Z_i > a\right) &= P\left(\exp\left(\sum_{i=1}^m \mu Z_i\right) > \exp(\mu a)\right), \\
&\leq \exp(-\mu a) \prod_{i=1}^m E(\exp(\mu Z_i)), \\
&\leq \exp(-\mu a) \prod_{i=1}^m (1 + \mu^2 E(Z_i^2)), \\
&\leq \exp(-\mu a) \exp\left(\mu^2 \sum_{i=1}^m E(Z_i^2)\right), \\
&\leq \exp(-\mu a) \exp(\mu^2 V).
\end{aligned}$$

On déduit le résultat en prenant $\mu = sa/V$. □

Lemme 3.10. Si $F \in \text{Lip}(1)$, définissons :

$$N_{x,m} = \{0 \leq k \leq m : |k - xm| \leq (m \log m)^{1/2}\}.$$

Alors pour $2 \leq m \leq (n \log n)^2$, on a $H_{n,m} = O(b_{n,m})$ p.s. quand $n \rightarrow \infty$

$$H_{n,m} = \sup_{0 < x < 1} \max_{k \in N_{x,m}} |F_n(k/m) - F(k/m) - F_n(x) + F(x)|. \quad (3.32)$$

Preuve. En subdivisant I en sous intervalles de longueur $b_{n,m}$, on peut trouver des entiers naturels i et j tels que : $jb_{n,m} < k/m \leq (j+1)b_{n,m}$ et $ib_{n,m} < x \leq (i+1)b_{n,m}$. En utilisant la croissance de F et F_n , on peut écrire :

$$\begin{aligned}
F_n(k/m) - F(k/m) - F_n(x) + F(x) &\leq F_n((j+1)b_{n,m}) - F(jb_{n,m}) - F_n(ib_{n,m}) + F((i+1)b_{n,m}), \\
&\leq F_n((j+1)b_{n,m}) - F_n(ib_{n,m}) - F((j+1)b_{n,m}) + F(ib_{n,m}), \\
&\quad + F((j+1)b_{n,m}) - F(jb_{n,m}) + F((i+1)b_{n,m}) - F(ib_{n,m}).
\end{aligned}$$

De même on a :

$$\begin{aligned}
F_n(k/m) - F(k/m) - F_n(x) + F(x) &\geq F_n(jb_{n,m}) - F((j+1)b_{n,m}) - F_n((i+1)b_{n,m}) + F(ib_{n,m}), \\
&\geq -\left[F_n((i+1)b_{n,m}) - F_n(jb_{n,m}) - F((i+1)b_{n,m}) + F(jb_{n,m}), \right. \\
&\quad \left. + F((i+1)b_{n,m}) - F(ib_{n,m}) + F((j+1)b_{n,m}) - F(jb_{n,m})\right].
\end{aligned}$$

Puisque $m \leq (n \log n)^2$ implique que $b_{n,m} \leq c_m$ et que $|x - k/m| \leq c_m$ pour $k \in N_{x,m}$, on peut écrire

$$H_{n,m} \leq D_{n,m} + D_{n,m,1}, \quad (3.33)$$

où

$$D_{n,m} = \max_{|i-j|b_{n,m} \leq 2c_m} |F_n((j+1)b_{n,m}) - F_n(ib_{n,m}) - F((j+1)b_{n,m}) + F(ib_{n,m})|,$$

et

$$D_{n,m,1} = \max_{i,j} |F(jb_{n,m}) - F((j+1)b_{n,m}) + F(ib_{n,m}) - F((i+1)b_{n,m})|.$$

Puisque $F \in \text{Lip}(1)$, il existe une constante $c \geq 2$ telle que :

$$|F(v) - F(u)| \leq c|u - v|,$$

on voit alors aisément que

$$D_{n,m,1} = O(b_{n,m}). \quad (3.34)$$

Montrons maintenant que $D_{n,m} = O(b_{n,m})$ p.s., $n \rightarrow \infty$. Pour $u \leq v$, posons :

$$Z_{i,u,v} = \mathbb{1}_{\{u \leq X_i \leq v\}} - (F(v) - F(u)).$$

D'après (3.2.1),

$$\sum_{k=0}^n \text{Var}(Z_{i,u,v}) = nE(Z_{1,u,v}^2) \leq n|F(v) - F(u)| \leq nc|u - v|.$$

Pour $|u - v| \leq 2c_m$, on peut appliquer le Lemme 3.9 avec $Z_i = Z_{i,u,v}$

(puis $Z_i = -Z_{i,u,v}$), $b = 1$, $s = 1/2$, $a = 4cnb_{n,m}$, et $V = 2cnc_m$ pour obtenir :

$$\begin{aligned} P(|F_n(u) - F_n(v) - F(u) + F(v)| > 4cb_{n,m}) &\leq 2\exp(-a^2/(4V)), \\ &\leq 2\exp(-2cnb_{n,m}^2/c_m), \\ &\leq 2\exp(-2c \log n), \\ &\leq 2n^{-2c}, \end{aligned}$$

puisque $\#\{(i,j) : 0 < ib_{n,m} \leq 1, 0 < jb_{n,m} \leq 1\} \leq b_{n,m}^{-2} \leq n^2$, on peut déduire que

$$P(D_{n,m} > 4cb_{n,m}) < 2n^2n^{-2c} \leq 2n^{-2},$$

et $\sum_{n \geq 0} P(D_{n,m} > 4cb_{n,m}) < \infty$. Le lemme de Borel-Cantelli nous permet de conclure que $D_{n,m} = O(b_{n,m})$ p.s., $n \rightarrow \infty$. On déduit le lemme de (3.33) et (3.34). \square

Preuve du Théorème 3.8.

Pour tout $x \in I$, on peut écrire

$$\tilde{F}_{n,m}(x) - F_n(x) = \sum_{k=0}^m [F_n(k/m) - F(k/m) - F_n(x) + F(x)] b_k(m, x) + (B_m(F)(x) - F(x)), \quad (3.35)$$

donc on a

$$\|\tilde{F}_{n,m} - F_n\| \leq \sup_{0 < x < 1} \left(\sum_{k=0}^m |D_{k,m,n}(x)| b_k(m, x) \right) + \|B_m(F) - F\|, \quad (3.36)$$

où

$$D_{k,m,n}(x) = F_n(k/m) - F(k/m) - F_n(x) + F(x).$$

Soit $0 \leq k \leq m$ fixé, puisque $f \in \text{Lip}(1)$ on a uniformément pour $x \in (0, 1)$,

$$F(k/m) - F(x) = (k/m - x)f(x) + O((k/m - x)^2).$$

En utilisant (1.1c) et (1.1b), on a rapidement

$$\|B_m(F) - F\| = O(m^{-1}). \quad (3.37)$$

Soit $Z_{i,x}$, $i = 1, 2, \dots, m$ des variables indépendantes de même loi, telle que $Z_{i,x} + x$ suive une loi de bernoulli de paramètre x . En appliquant le Lemme 3.9, avec $Z_i = Z_{i,x}$ (puis $Z_i = -Z_{i,x}$), $a = (m \log m)^{1/2}$, $V = m/4$ et $s = 1/2$, pour $m \geq 16$, on a :

$$\sum_{k \notin N_{x,m}} b_k(m, x) \leq 2 \exp(-4m \log m / (4m)) = 2m^{-1}. \quad (3.38)$$

Puisque $|D_{k,m,n}(x)| \leq 2$ pour tout k, m, n et x , on peut écrire en utilisant (3.38) et le Lemme 3.10 que

$$\begin{aligned} \sum_{k=0}^m |D_{k,m,n}(x)| b_k(m, x) &\leq \sum_{k \in N_{x,m}} |D_{k,m,n}(x)| b_k(m, x) + \sum_{k \notin N_{x,m}} |D_{k,m,n}(x)| b_k(m, x), \\ &\leq H_{n,m} \sum_{k=0}^m b_k(m, x) + 2 \sum_{k \notin N_{x,m}} b_k(m, x), \\ &\leq H_{n,m} + 4m^{-1}. \end{aligned} \quad (3.39)$$

Puisque $m^{-1} \leq b_{n,m}$ pour $n^{2/3} \leq m$, les équations (3.36), (3.37) et le Lemme 3.10 nous permettent de conclure. \square

En reprenant les notations (3.8), (3.9), (3.10) et vu (3.11) un estimateur naturel de la densité f est

$$\begin{aligned}\tilde{f}_{n,m}(x) &= \sum_{k=0}^m f_n(k/m) \frac{d}{dx} \mathcal{B}_k(m, x), \\ &= m \sum_{k=0}^{m-1} f_n((k+1)/m) b_k(m-1, x).\end{aligned}\quad (3.40)$$

Cet estimateur (à une constante près) fut proposé pour la première fois par Vitale (1975) qui démontra ces premières propriétés. Le résultat suivant nous montre que sous certaines hypothèses $\tilde{f}_{n,m}$ est un bon estimateur de f .

Théorème 3.11. *Pour $2 \leq m \leq n/\log n$, on a p.s. pour $n \rightarrow \infty$,*

$$\|\tilde{f}_{n,m} - f\| = O(m^{1/2} c_n) + O(\|B'_m(F) - f\|), \quad (3.41)$$

donc si $m = o(n/\log n)$, alors $\|\tilde{f}_{n,m} - f\| \rightarrow 0$ p.s. quand $n, m \rightarrow \infty$.

Preuve. On peut écrire

$$\begin{aligned}\tilde{f}_{n,m}(x) &= m \sum_{k=0}^{m-1} (F_n((k+1)/m) - F_n(k/m)) b_k(m-1, x), \\ &= m \sum_{k=0}^{m-1} (F((k+1)/m) - F(k/m)) b_k(m-1, x) \\ &\quad + m \sum_{k=0}^{m-1} (F_n((k+1)/m) - F((k+1)/m) - F_n(k/m) + F(k/m)) b_k(m-1, x), \\ &= T_m(x) + T_{m,n}(x),\end{aligned}\quad (3.42)$$

où

$$T_m(x) = m \sum_{k=0}^{m-1} (F((k+1)/m) - F(k/m)) b_k(m-1, x), \quad (3.43)$$

et

$$T_{m,n}(x) = m \sum_{k=0}^{m-1} (F_n((k+1)/m) - F((k+1)/m) - F_n(k/m) + F(k/m)) b_k(m-1, x). \quad (3.44)$$

Il existe une constance $c > 2$ telle que $F((k+1)/m) - F(k/m) \leq c/m$. Soit

$$Z_{i,k} = \mathbb{1}_{k \leq mX_i \leq k+1} - (F((k+1)/m) - F(k/m)).$$

On peut appliquer le Lemme 3.9 aux variables aléatoires $Z_i = Z_{i,k}$ (puis $Z_i = -Z_{i,k}$), $b = 1$, $s = 1/2$, $a = 2cnc_n m^{-1/2}$ et $V = cn/m$ pour avoir

$$P(L_{m,n} > 2cc_n m^{-1/2}) \leq 2me^{-a^2/(4V)} = 2m/n^c \leq 2n^{1-c}, \quad (3.45)$$

où

$$L_{m,n} = \sup_{0 \leq k \leq m-1} |F_n((k+1)/m) - F((k+1)/m) - F_n(k/m) + F(k/m)|.$$

On a utilisé la condition $m \leq n/\log n$ pour avoir $a \leq 2V$. Lemme de Borel-Cantelli nous permet de conclure que $L_{m,n} = O(m^{-1/2}c_n)$ p.s. quand $n \rightarrow \infty$ et finalement $\|T_{m,n}\| = O(m^{1/2}c_n)$ p.s. quand $n \rightarrow \infty$. On écrit ensuite en remarquant que

$$T_m = B'_m(F),$$

$$\|\tilde{f}_{n,m} - f\| = \|T_m + T_{m,n} - f\| \leq \|T_m - f\| + \|T_{m,n}\| \leq O(m^{1/2}c_n) + O(\|B'_m(F) - f\|).$$

Puisque f est supposé de Lipschitz d'ordre 1, en utilisant la formule de Taylor et (1.8), on a uniformément pour $0 \leq x \leq 1$ que :

$$\begin{aligned} B'_m(F)(x) - f(x) &= \sum_{k=0}^{m-1} (f(k/m) - f(x))b_k(m-1, x) + O(m^{-1}), \\ &= O\left(\sum_{k=0}^{m-1} |k/m - x|b_k(m-1, x)\right) + O(m^{-1}), \\ &= O(m^{-1/2}). \end{aligned} \quad (3.46)$$

□

3.2.2. Loi du logarithme itéré

Le but de cette section est de montrer que l'estimateur de Bernstein $\tilde{F}_{n,m}$, comme l'estimateur par noyau (voir Winter (1979)) vérifie (3.6) sous certaines conditions de régularité.

Théorème 3.12. (Voir Leblanc (2009)). Si $F \in \text{Lip}(\alpha)$ et si $ma_n^{-2/\alpha} \rightarrow \infty$ pour $n, m \rightarrow \infty$ alors

$$\limsup_{n,m \rightarrow \infty} a_n \|F_{n,m} - F\| \leq 1 \quad p.s. \quad (3.47)$$

Preuve. Puisque $F \in \text{Lip}(\alpha)$, on voit aisément qu'il existe $c > 0$ tel que $\omega_F(\delta) \leq c\delta^\alpha$, pour tout $\delta > 0$. On peut alors déduire de (1.6) que :

$$\|B_m(F) - F\| \leq \frac{3}{2}cm^{-\alpha/2}. \quad (3.48)$$

En partant de (3.28) et vu (3.29) on peut écrire :

$$\begin{aligned} \limsup_{n,m \rightarrow \infty} a_n \|F_{n,m} - F\| &\leq \limsup_{n \rightarrow \infty} a_n \|F_n - F\| + \limsup_{n,m \rightarrow \infty} a_n \|B_m(F) - F\|, \\ &= 1 + \limsup_{n,m \rightarrow \infty} a_n m^{-\alpha/2} = 1. \end{aligned}$$

Le résultat en découle. \square

Théorème 3.13. (Voir Leblanc (2009)). Si $f \in \text{Lip}(\alpha)$ et si $ma_n^{-2/(1+\alpha)} \rightarrow \infty$ pour $n, m \rightarrow \infty$ alors

$$\limsup_{n,m \rightarrow \infty} a_n \|F_{n,m} - F\| \leq 1 \quad p.s. \quad (3.49)$$

Preuve. Il existe une constante $c > 0$ telle que $\omega_f(\delta) \leq c\delta^\alpha$, l'équation (1.9) nous permet alors d'écrire

$$\|B_m(F) - F\| \leq \frac{3}{4}cm^{-(1+\alpha)/2}. \quad (3.50)$$

On conclut comme dans la preuve du Théorème 3.12. \square

Soit Γ_n la fonction de répartition empirique associée à un échantillon $\xi_1, \xi_2, \dots, \xi_n$ de loi uniforme sur $[0,1]$. Le processus empirique associé à Γ_n s'écrit

$$\mathbb{U}_n(t) = n^{1/2}(\Gamma_n(t) - t). \quad (3.51)$$

Le module d'oscillation est

$$\omega_n(u) = \sup_{0 \leq h \leq u} \sup_{0 \leq t \leq 1-h} |\mathbb{U}_n(t+h) - \mathbb{U}_n(t)|. \quad (3.52)$$

Les deux lemmes suivants sont très importants pour la suite.

Lemme 3.14. Si $F \in \text{Lip}(\alpha)$, alors il existe une constante c telle que pour $m, n \rightarrow \infty$,

$$\|\tilde{F}_{n,m} - F_n\| = n^{-1/2}\omega_n(cb_m) + O(m^{-\alpha/2}), \quad (3.53)$$

où $b_m = (m^{-1} \log m)^{\alpha/2}$.

Preuve. D'après (3.36) et (3.39), on peut écrire :

$$\|\tilde{F}_{n,m} - F_n\| \leq H_{n,m} + 4m^{-1} + \|B_m(F) - F\|.$$

Pour $k \in N_{x,m}$ et $x \in I$, puisque $F \in \text{Lip}(\alpha)$, il existe une constance $c > 0$ tel que

$$|F(k/m) - F(x)| \leq c|k/m - x|^\alpha \leq c(m^{-1} \log m)^{\alpha/2} = cb_m. \quad (3.54)$$

Puisque $b_m \rightarrow 0$, $m \rightarrow \infty$, pour m assez grand on a $cb_m \leq 1$. Le module de continuité étant croissant, vu (3.1) et (3.54) on peut écrire

$$\begin{aligned} n^{1/2}|D_{k,m,n}(x)| &= n^{1/2}|\Gamma_n(F(k/m)) - F(k/m) - \Gamma_n(F(x)) - F(x)|, \\ &= |U_n(F(k/m)) - U_n(F(x))|, \\ &\leq \omega_n(|F(k/m) - F(x)|), \\ &\leq \omega_n(cb_m). \end{aligned}$$

On déduit alors que :

$$\|\tilde{F}_{n,m} - F_n\| \leq n^{-1/2}\omega_n(cb_m) + 4m^{-1} + \|B_m(F) - F\|.$$

Le résultat découle alors de (3.48). \square

Lemme 3.15. Si $f \in \text{Lip}(\alpha)$, alors il existe une constante c telle que pour $m, n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = n^{-1/2}\omega_n(\|f\|b_m) + O(m^{-(1+\alpha)/2}), \quad (3.55)$$

où $b_m = (m^{-1} \log m)^{1/2}$.

Preuve. La preuve est similaire à celle du lemme précédent en remarquant que $f \in \text{Lip}(\alpha)$ implique $F \in \text{Lip}(1)$ et en appliquant (3.50) au lieu de (3.48). \square

Soit b_m une suite d'éléments de I vérifiant $b_m \searrow 0$. Faisons les hypothèses suivantes :

- Hypothèse 1 $\log(1/b_m)/\log \log n \rightarrow a \in [0, \infty)$ et $nb_m \nearrow$ quand $n \rightarrow \infty$.
- Hypothèse 2 $\log(1/b_m)/\log \log n \rightarrow \infty$, $nb_m/\log n \rightarrow \infty$ et $nb_m \nearrow$ quand $n \rightarrow \infty$.
- Hypothèse 3 $nb_m/\log n \rightarrow 0$, $\log(1/b_m)/\log n \rightarrow 1$ et $n^2b_m \nearrow$ quand $n \rightarrow \infty$.
- Hypothèse 4 $\log(1/b_m)/\log n = c \in (1, \infty]$.

On a les théorèmes suivants

Théorème 3.16. (Voir Leblanc (2009)).

Soit $F \in Lip(\alpha)$. Pour $b_m = (m^{-1} \log m)^{\alpha/2}$, on a les résultats suivants :

i) si l'hypothèse 1 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(m^{-\alpha/2}),$$

ii) si l'hypothèse 2 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(m^{-\alpha/2}),$$

iii) si l'hypothèse 3 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O((n \log \log n)^{-1} \log n) + O(m^{-\alpha/2}),$$

iv) si l'hypothèse 4 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(n^{-1}).$$

Preuve. Dans le cas i), une application du Théorème 2.1 (a) et du Lemme 3.14, nous donne

$$\|\tilde{F}_{n,m} - F_n\| = O(n^{-1/2}(\log \log n)^{1/2} b_m^{1/2}) + O(m^{-\alpha/2}), \text{ p.s.} \quad (3.56)$$

Leblanc (2009) affirme que sous les hypothèses i), le second terme du membre de droite de (3.56) domine le premier terme. Les cas ii), iii) et iv) se prouvent de manière similaire. On applique le Théorème 2.1. \square

Théorème 3.17. Soit $f \in Lip(\alpha)$. Pour $b_m = (m^{-1} \log m)^{1/2}$, on a les résultats suivants :

i) si l'hypothèse 1 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(m^{-(1+\alpha)/2}),$$

ii) si l'hypothèse 2 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(n^{-1/2} m^{-1/4} (\log m)^{3/4}) + O(m^{-(1+\alpha)/2}),$$

iii) si l'hypothèse 3 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O((n \log \log n)^{-1} \log n),$$

iv) si l'hypothèse 4 est vérifiée, alors pour $n \rightarrow \infty$

$$\|\tilde{F}_{n,m} - F_n\| = O(n^{-1}).$$

Preuve. La preuve est similaire à celle du Théorème 3.16. On applique le Théorème 2.1 et le Lemme 3.15. \square

Nous pouvons enfin montrer que l'estimateur de Bernstein $\tilde{F}_{n,m}$ vérifie la relation (3.6) sous certaines conditions.

Théorème 3.18. Soit $F \in Lip(\alpha)$. Si $ma_n^{-2/\alpha} \rightarrow \infty$, $m, n \rightarrow \infty$ alors sous les hypothèses 2, 3 ou 4, l'estimateur $\tilde{F}_{n,m}$ vérifie la relation (3.6).

Preuve. D'après le Théorème 3.13, il suffit de montrer que

$$\limsup_{n \rightarrow \infty} a_n \|\tilde{F}_{n,m} - F\| \geq 1 \quad \text{p.s.}$$

L'inégalité triangulaire donne

$$\|\tilde{F}_{n,m} - F\| \geq \|F_n - F\| - \|\tilde{F}_{n,m} - F_n\|,$$

on a immédiatement

$$\limsup_{n \rightarrow \infty} a_n \|\tilde{F}_{n,m} - F\| \geq \limsup_{n \rightarrow \infty} a_n \|F_n - F\| - \limsup_{n \rightarrow \infty} a_n \|\tilde{F}_{n,m} - F_n\|.$$

Les Théorèmes 3.3 et 3.16 permettent de conclure. \square

Théorème 3.19. Soit $f \in Lip(\alpha)$. Si $ma_n^{-2/(1+\alpha)} \rightarrow \infty$, $m, n \rightarrow \infty$ alors sous les hypothèses 2, 3 ou 4, l'estimateur $\tilde{F}_{n,m}$ vérifie la relation (3.6).

Preuve. La preuve est similaire à celle du Théorème 3.18. On utilise les Théorèmes 3.8, 3.3 et 3.17. \square

3.2.3. Résultats asymptotiques

Nous présentons ici quelques résultats asymptotiques concernant les estimateurs de Bernstein. Ces résultats sont tirés de Babu et al. (2002) et Leblanc (2012a). Commençons par ceux concernant la fonction de repartitions $\tilde{F}_{n,m}$.

Théorème 3.20. (Voir Babu et al. (2002)). Si F est deux fois continûment dérivable à dérivées bornées sur I , vérifiant $0 < F(x) < 1$ pour $x \in (0, 1)$, lorsque $n, m \rightarrow \infty$ alors on a :

$$n^{1/2}(\tilde{F}_{n,m}(x) - B_m(F)(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad (3.57)$$

donc si $mn^{-1/2} \rightarrow \infty$

$$n^{1/2}(\tilde{F}_{n,m}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad (3.58)$$

et si $mn^{-1/2} \rightarrow c, c > 0$,

$$n^{1/2}(\tilde{F}_{n,m}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(c^{-1}b_1(x), \sigma^2(x)). \quad (3.59)$$

Preuve. Commençons par prouver (3.57). Reprenons la définition des variables aléatoires $Y_{i,m}$ $i = 1, 2, \dots, n$ et $k = 0, 1, \dots, m$ comme en (3.16). On a,

$$n^{1/2}(\tilde{F}_{n,m}(x) - B_m(x)) = n^{-1/2} \sum_{i=1}^n Y_{i,m}$$

L'idée est d'appliquer le théorème de Lindeberg sur les variables aléatoires i.i.d. $Y_{i,m}$ $i = 1, 2, \dots, n$. On a $s_{n,m}^2 = n\text{Var}(Y_{1,m}) = nE(Y_{1,m}^2) = n\sigma^2(x)$ et $|Y_{1,m}| \leq 2 \sum_{k=0}^m b_k(m, x) = 2$. On voit que pour tout $\varepsilon > 0$,

$$s_{n,m}^{-2} \sum_{i=1}^n E(Y_{i,m}^2 \mathbb{1}_{\{|Y_{i,m}| > \varepsilon s_{n,m}\}}) \rightarrow 0 \quad (n \rightarrow \infty),$$

Le Théorème A.2 nous permet alors de déduire (3.57).

En utilisant (1.11) avec $u = F$, on a rapidement

$$n^{1/2}(\tilde{F}_{n,m}(x) - F(x)) = n^{1/2}(\tilde{F}_{n,m}(x) - B_m(F)(x)) + m^{-1}n^{1/2}b_1(x) + o(m^{-1}n^{1/2}).$$

Le lemme de Slutsky pour permet alors de déduire aisément (3.58) et (3.59). □

Posons

$$\gamma(x) = f(x)(4\pi x(1-x))^{-1/2} = f(x)\gamma_1(x), \quad 0 < x < 1.$$

Reprenons l'expression $\tilde{f}_{n,m} = T_m + T_{m,n}$ avec les définitions de T_m et $T_{m,n}$ comme en (3.43) et (3.44), on a le résultat suivant.

Proposition 3.21. Si $f(x) > 0$, alors

$$n^{1/2}m^{-1/4}T_{m,n}(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma(x)), \quad (3.60)$$

quand $n, m \rightarrow \infty$ avec $2 \leq m \leq n/\log n$ et $mn^{-1} \rightarrow 0$.

Par souci de simplification posons

$$a_{k,m} = F((k+1)/m) - F(k/m), \quad (3.61)$$

et

$$Y_{i,m} = \sum_{k=0}^{m-1} (\mathbb{1}_{\{k \leq mX_i \leq k+1\}} - a_{k,m}) b_k(m-1, x). \quad (3.62)$$

On voit alors que

$$T_{m,n}(x) = \frac{m}{n} \sum_{i=1}^n Y_{i,m}. \quad (3.63)$$

Puisque les $Y_{i,m}$ sont des variables aléatoires i.i.d., on a

$$\text{Var}(T_{m,n}(x)) = m^2 n^{-1} \text{Var}(Y_{1,m}).$$

Afin de prouver la Proposition 3.21, on a besoin du lemme suivant :

Lemme 3.22. *Si $f(x) > 0$, alors*

$$m^{3/2} \text{Var}(Y_{1,m}) \rightarrow \gamma(x), \quad m \rightarrow \infty$$

Preuve.

On voit rapidement de (3.62) que

$$\text{Var}(Y_{1,m}) = \sum_{k=0}^{m-1} b_k^2(m-1, x) a_{k,m} - \left(\sum_{k=0}^{m-1} b_k(m-1, x) a_{k,m} \right)^2. \quad (3.64)$$

En appliquant deux fois la formule de Taylor on a :

$$\begin{aligned} m a_{k,m} &= f(k/m) + O(m^{-1}), \\ &= f(x) + O(|x - k/m|) + O(m^{-1}). \end{aligned} \quad (3.65)$$

Puisque $b_k(m, x) \leq 1$, en utilisant l'inégalité de Cauchy-Schwarz, (1.1c) et (i) du Lemme 3.10, on a

$$\begin{aligned}
 \left(\sum_{k=0}^{m-1} b_k^2(m, x) |x - k/m| \right)^2 &\leq \left(\sum_{k=0}^{m-1} b_k^3(m, x) \right) \left(\sum_{k=0}^{m-1} b_k(m, x) (x - k/m)^2 \right), \\
 &\leq O \left(m^{-1} \sum_{k=0}^{m-1} b_k^3(m, x) \right), \\
 &= O \left(m^{-1} \sum_{k=0}^{m-1} b_k^2(m, x) \right), \\
 &= O(m^{-3/2}).
 \end{aligned} \tag{3.66}$$

En utilisant (3.65), (3.66) et une nouvelle application de (i) du Lemme 3.10, on a

$$\begin{aligned}
 m \sum_{k=0}^{m-1} b_k^2(m, x) a_{k,m} &= f(x) \sum_{k=0}^{m-1} b_k^2(m, x) + O(m^{-1}) + O \left(\sum_{k=0}^{m-1} b_k^2(m, x) |x - k/m| \right), \\
 &= m^{-1/2} \gamma(x) (1 + o(1)) + O(m^{-3/4}).
 \end{aligned}$$

Puisque d'après (3.46)

$$m \sum_{k=0}^{m-1} a_{k,m} b_k(m-1, x) = T_m(x) = f(x) + O(m^{-1/2}) = O(1),$$

on déduit des relations (3.64) et (3.66) que

$$m^{3/2} \text{Var}(Y_{1,m}) = \gamma(x) (1 + o(1)) + O(m^{-1/4}) - O(m^{-2}) = \gamma(x) (1 + o(1)) + O(m^{-1/4}).$$

Et le résultat en découle. \square

Preuve de Proposition 3.21. Nous allons appliquer le théorème de Lindeberg aux variables aléatoires i.i.d., $Y_{i,m}$, $i = 1, 2, \dots, n$. Montrons donc que pour tout $\varepsilon > 0$,

$$s_{n,m}^{-2} \sum_{i=1}^n E(Y_{i,m}^2 I_{\{|Y_{i,m}| > \varepsilon s_{n,m}\}}) \rightarrow 0 \quad (n \rightarrow \infty), \tag{3.67}$$

où $s_{n,m}^2 = n \text{Var}(Y_{1,m})$.

D'après le théorème des accroissements finis, il existe $k/m \leq \xi \leq (k+1)/m$ tel que

$$m a_{k,m} = f(\xi). \tag{3.68}$$

Soit $\omega \in \Omega$, il existe $0 \leq j \leq m-1$ tel que $j \leq mX_i(\omega) \leq j+1$, donc

$$\sum_{k=0}^{m-1} \mathbb{1}_{\{k \leq mX_i(\omega) \leq k+1\}} b_k(m-1, x) = b_j(m-1, x),$$

en utilisant (i) du Lemme 3.10 et (3.68) on a alors

$$\begin{aligned} |Y_{i,m}| &\leq \max_{0 \leq k \leq m-1} a_{k,m} + \max_{0 \leq k \leq m-1} b_k(m-1, x), \\ &= O(m^{-1}) + \left(\sum_{k=0}^{m-1} b_k^2(m-1, x) \right)^{1/2}, \\ &= O(m^{-1/4}). \end{aligned}$$

D'après le Lemme 3.22, $s_{n,m}^2 = nm^{-3/2}\gamma(x)(1 + o(1))$, ce qui donne

$$|Y_{i,m}| s_{n,m}^{-1} = O(m^{-1/4} m^{3/4} n^{-1/2}) = O((mn^{-1})^{1/2}) \rightarrow 0, \quad n, m \rightarrow \infty,$$

car par hypothèse $mn^{-1} \rightarrow 0$. On voit alors que la condition (3.67) est vérifiée. Le théorème de Lindeberg (Théorème A.2) nous permet de conclure. \square

On peut enfin prouver des résultats montrant la normalité asymptotique de l'estimateur $\tilde{f}_{n,m}$.

Théorème 3.23. (Voir Babu et al. (2002)). Si $f(x) > 0$, alors

$$n^{1/2} m^{-1/4} (\tilde{f}_{n,m} - f)(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma(x)), \quad (3.69)$$

quand $n, m \rightarrow \infty$ avec $2 \leq m \leq n/\log n$ et $n^{2/3} m^{-1} \rightarrow 0$.

Théorème 3.24. (Voir Babu et al. (2002)). On suppose que f est deux fois continûment dérivable sur un voisinage de $x \in (0, 1)$. Si $f(x) > 0$ et si pour $n, m \rightarrow \infty$, $n^{-1/5} m^{1/2} \rightarrow \delta$, pour $\delta > 0$, alors

$$n^{2/5} (\tilde{f}_{n,m} - f)(x) - \frac{1}{2} \delta^{-2} b(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma(x) \delta), \quad (3.70)$$

où $b(x) = (1 - 2x)f'(x) + x(1 - x)f''(x)$

Preuve du Théorème 3.23. On peut écrire

$$n^{1/2} m^{-1/4} (\tilde{f}_{n,m} - f) = n^{1/2} m^{-1/4} (T_m - f) + n^{1/2} m^{-1/4} T_{m,n}$$

D'après le lemme de Slutsky et la Proposition 3.21 (la condition $2 \leq m \leq n/\log n$ nous garantit que $mn^{-1} \rightarrow 0$) il suffit pour conclure de montrer que

$$n^{1/2}m^{-1/4}(T_m(x) - f(x)) = o(1).$$

Mais d'après (3.46), $m^{1/2}(T_m(x) - f(x)) = O(1)$ et $n^{1/2}m^{-3/4} = (n^{2/3}m^{-1})^{3/4}$. On déduit alors le résultat de l'hypothèse $n^{2/3}m^{-1} \rightarrow 0$. \square

Preuve du Théorème 3.24. On écrit encore

$$n^{2/5}(\tilde{f}_{n,m} - f)(x) = n^{2/5}(T_m - f) + n^{2/5}T_{m,n}.$$

On s'intéresse premièrement à $n^{2/5}(T_m - f)$. Soit $k \in N_{x,m}$. On peut écrire d'après la formule de Taylor :

$$\begin{aligned} m a_{k,m} &= f(k/m) + \frac{1}{2m}f'(k/m) + O(m^{-2}), \\ &= f(x) + \frac{1}{2m}f'(x) + o(m^{-1}), \end{aligned} \quad (3.71)$$

et

$$\begin{aligned} m(f(k/m) - f(x)) &= (k - xm)f'(x) + \frac{1}{2}m^{-1}(k - xm)^2f''(x)(1 + o(1)), \\ &= (k - (m-1)x)f'(x) - xf'(x) \\ &\quad + \frac{1}{2}m^{-1}(k - xm)^2f''(x)(1 + o(1)). \end{aligned} \quad (3.72)$$

On déduit de (3.71) et (3.72) que

$$2m(T_m(x) - f(x)) = b(x) + o(1), \quad (3.73)$$

donc

$$n^{2/5}(T_m(x) - f(x)) = \frac{1}{2m}n^{2/5}b(x) + o(n^{2/5}m^{-1}),$$

et par conséquent

$$n^{2/5}(T_m(x) - f(x)) = \frac{1}{2\delta^2}b(x) + o(1),$$

on déduit alors le résultat de la Proposition 3.21, car $mn^{-1}n = (m^{1/2}n^{-1/5})^2n^{-3/5} \rightarrow 0$ et $n^{2/5}T_{m,n} = (m^{1/2}n^{-1/5})^{1/2}n^{1/2}m^{-1/4}T_{m,n} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma(x)\delta)$. \square

Chapitre 4

ETUDES NUMÉRIQUES

Dans ce chapitre, nous présentons des résultats de simulations (obtenus à partir du logiciel R) afin d'apprécier le comportement des estimateurs de Bernstein.

4.1. FONCTION DE RÉPARTITION

Nous nous sommes intéressés à 3 distributions sur I : (i) $\text{Beta}(3,2)$ qui a une fonction de répartition à peu près symétrique et présentant une courbure très faible (ii) $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$ qui présente une courbure plus importante que la distribution $\text{Beta}(3,2)$ et enfin (iii) $1/2\text{Beta}(12,1)+1/2\text{Beta}(9,1)$ qui est complètement asymétrique. Les Figures 4.1, 4.2 et 4.3 nous présentent le comportement des estimateurs de Bernstein pour 10 simulations de taille $n = 25, 75$ et 100 . Pour chacune des illustrations, nous avons fait les choix $m = n^{2/3}$ et $m = n$. D'après les théorèmes du chapitre 3, le choix de m est très important pour avoir une bonne estimation. Si on ne s'intéresse qu'à une estimation ponctuelle ou sur un segment de I , il pourrait être important de savoir quelle est la valeur de m donnant la meilleure approximation de $F(x)$ pour un x fixé. La valeur optimale de $m = m_{\text{opt}}(x)$, a été obtenue par simulation de Monte-Carlo, en chaque point x de discrétisation de l'intervalle I . Nous avons déterminé la valeur de m , $2 \leq m \leq 200$ minimisant l'EQM que nous avons approchée par

$$\frac{1}{M} \sum_{i=1}^M [\tilde{F}_{n,m}^i(x) - F(x)]^2,$$

où $\tilde{F}_{n,m}^i(x)$ est la i -ème estimation de $\tilde{F}_{n,m}(x)$ basée sur un échantillon de taille n . Nous avons effectué une discrétisation uniforme de pas $h = 1/100$ de I et avons

choisi $M = 1000$ pour la simulation de Monte-Carlo. Ce choix se justifie par le nombre important d'itérations machine. Par exemple pour $M = 10000$ nos simulations auraient mis approximativement 1 mois et 22 jours. (Évaluation faite sur un ordinateur avec un processeur quadruple cœurs cadencés à 3.20 GHz et une mémoire vive de 6 GB.) Notons que les mêmes simulations ont été utilisées en chaque point de discrétisation, nous n'avons considéré ici que les cas $n = 25$ et $n = 100$.

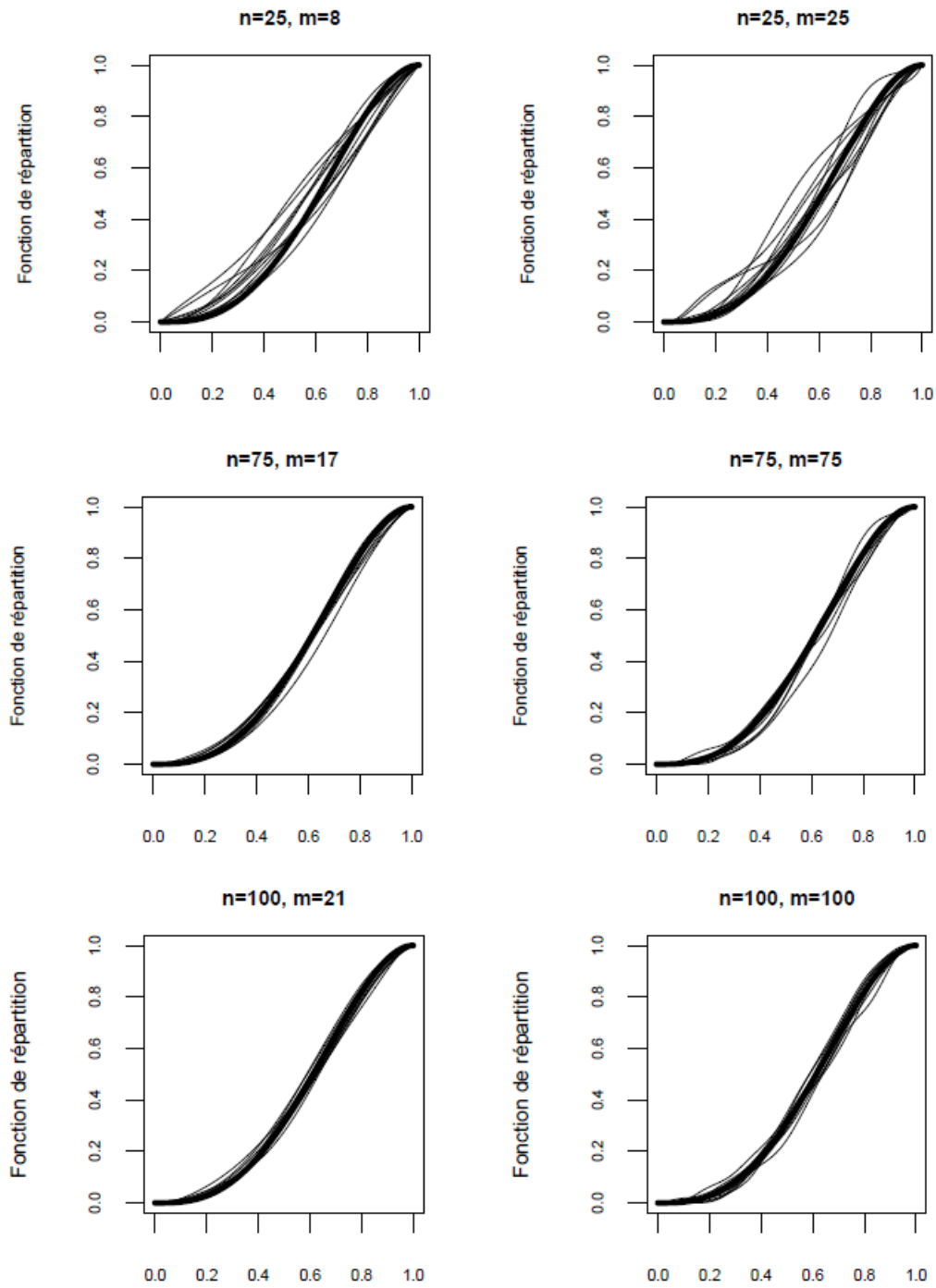


FIGURE 4.1. Estimateur de Bernstein pour un échantillon de loi Beta(3,2).

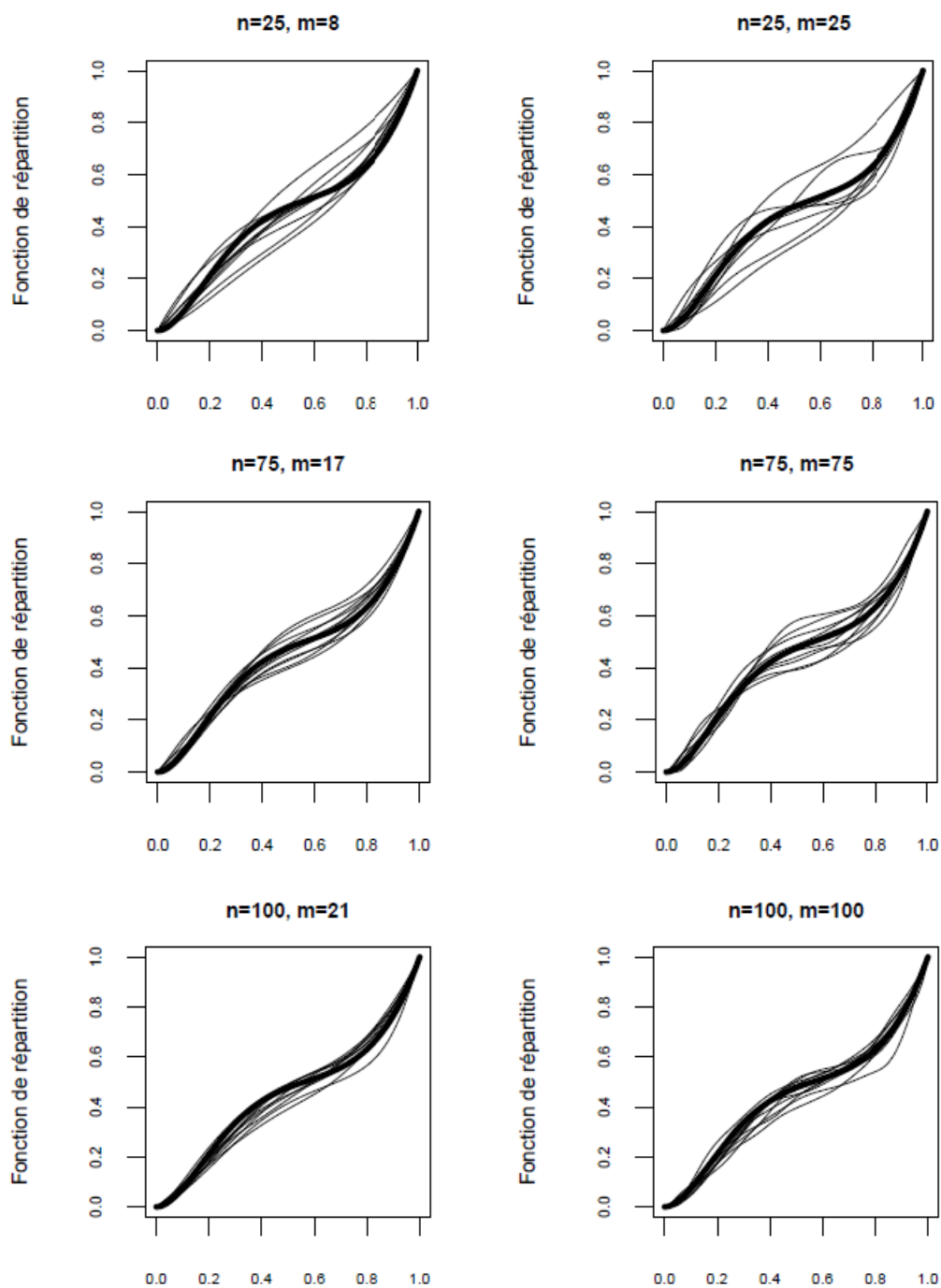


FIGURE 4.2. Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(1,6)$.

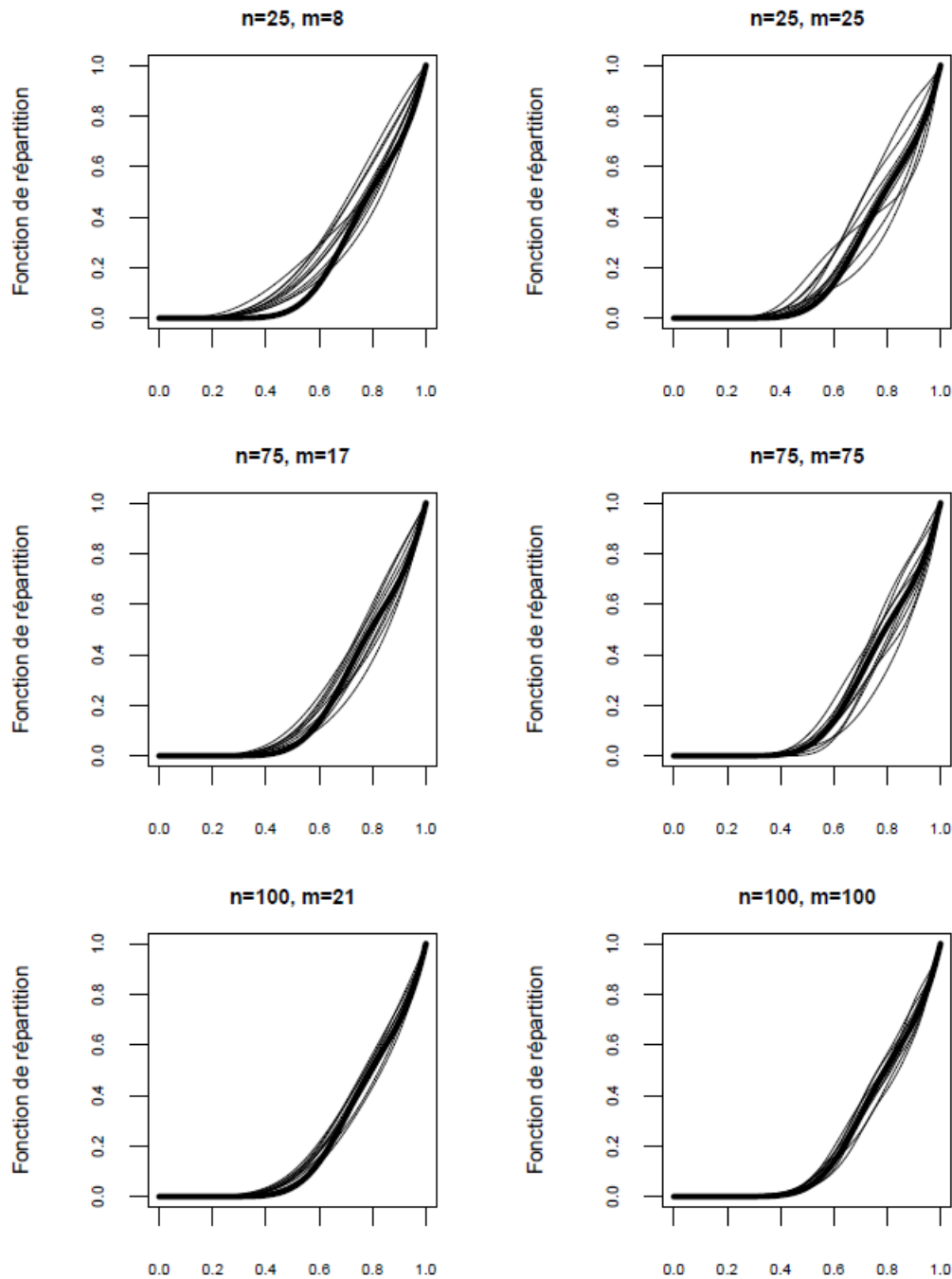


FIGURE 4.3. Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$.

Les Figures 4.4, 4.6 et 4.8 nous présentent les valeurs de l'EQM obtenues. On peut remarquer que dans les trois cas, l'EQM pour le choix optimal de m est meilleure que celle de la fonction de répartition empirique. On a aussi la confirmation que le cas $m = n^{2/3}$ est globalement meilleur (dans le sens de l'EQMI qui peut être approchée par l'aire sous chacune des courbes) pour les cas (i) et (ii) alors que $m = n$ est globalement meilleur pour le cas (iii). Les figures 4.5, 4.7 et 4.9 nous présentent les valeurs de m choisies à l'optimum, on note dans le cas (i) et très particulièrement dans le cas (iii) que pour certaines valeurs de x comprise entre 0 et 0.2 la valeur maximale $m = 200$ est atteinte ce qui peut laisser penser que le véritable maximum est au-delà de 200, mais une étude similaire, que nous ne présentons pas ici, avec $m = 500$ montre que l'EQM est très petit autour de ces valeurs et qu'un choix de m plus petit suffit pour avoir une estimation acceptable. Le choix asymptotique donné au Corollaire 3.6 pourrait suffire à avoir une bonne estimation. Pour vérifier cela nous avons effectué 10 simulations pour chacune des distributions avec à chaque point x de discrétisation les choix des m optimaux (m_{opt}) et asymptotiques (m_{aopt}). Les Figures 4.10, 4.11 et 4.12 nous montrent que pour n grand l'approximation est bonne en dehors des voisinages des points de discrétisation tels que la dérivée de f en x est nulle.

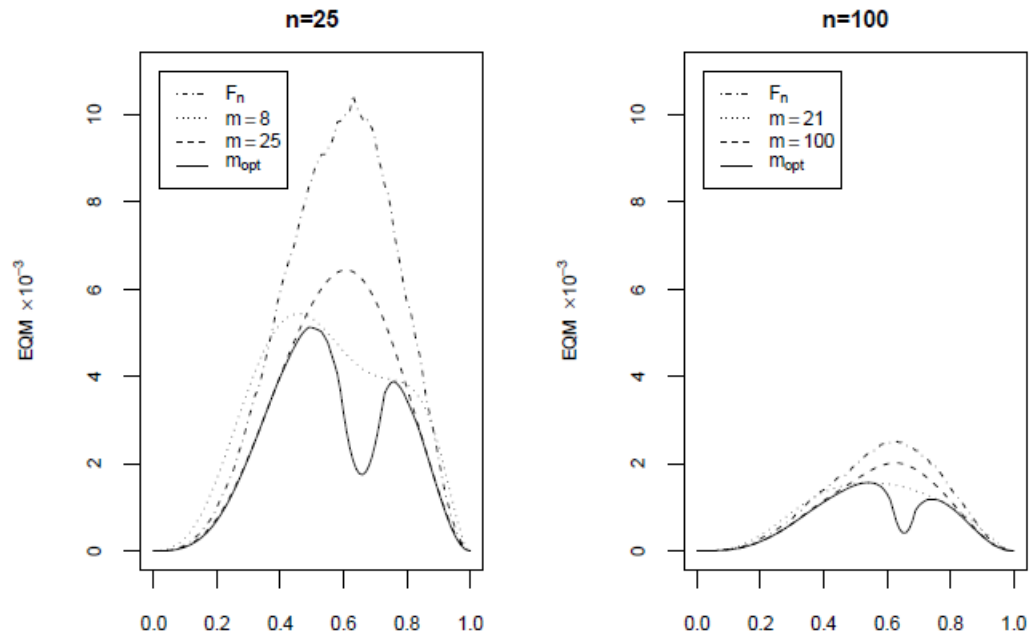


FIGURE 4.4. Erreur d'estimation pour un échantillon de loi Beta(3,2).

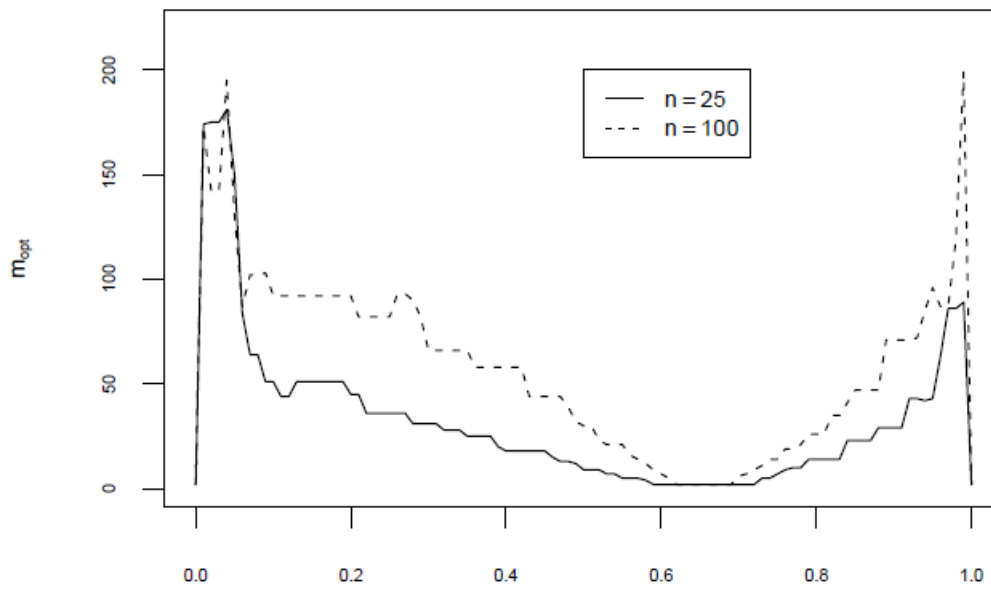


FIGURE 4.5. Valeur optimale de m pour un échantillon de loi Beta(3,2).

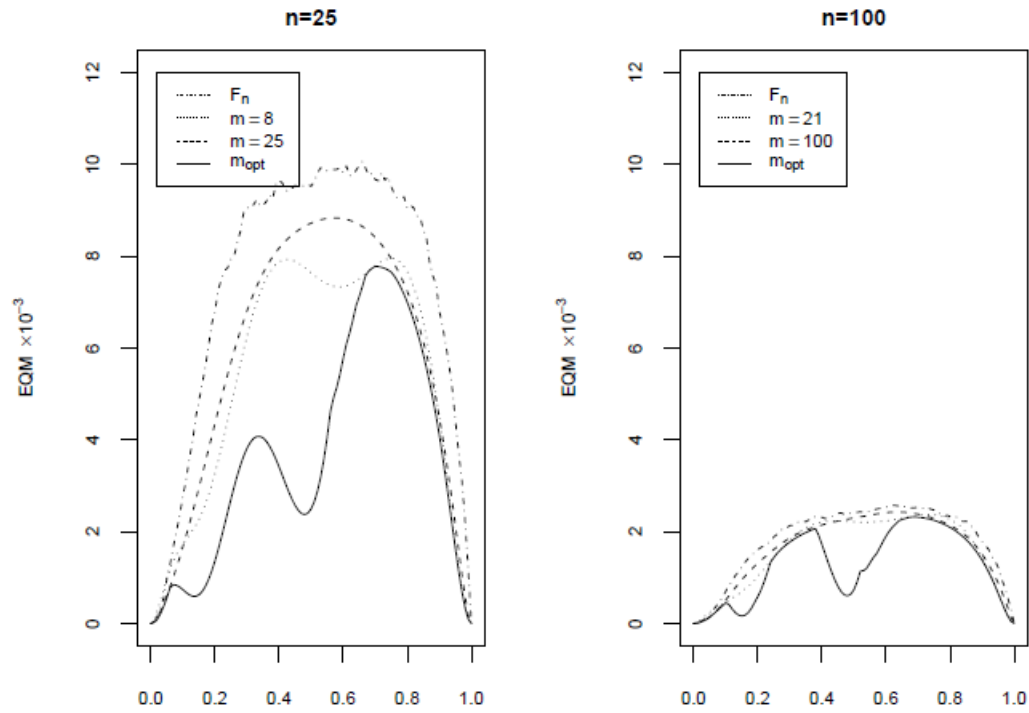


FIGURE 4.6. Erreur d'estimation pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$.

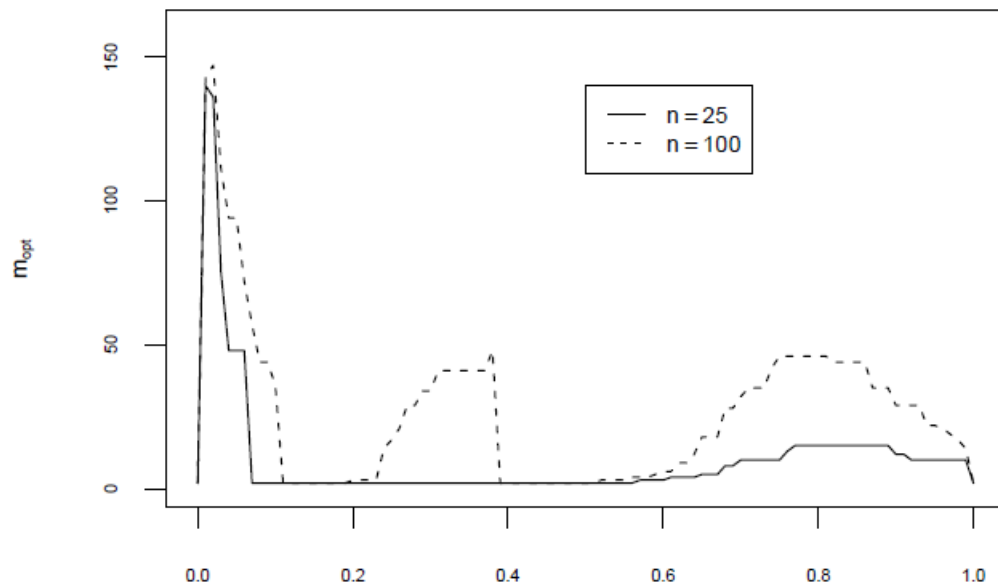


FIGURE 4.7. Valeur optimale de m pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$.

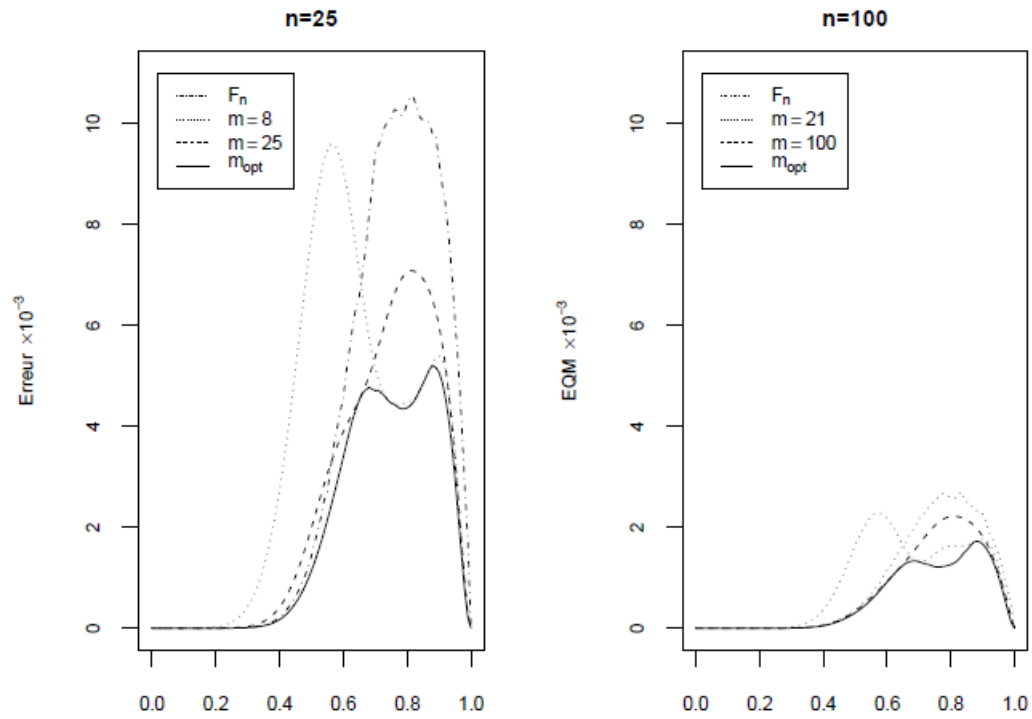


FIGURE 4.8. Erreur d'estimation pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$.

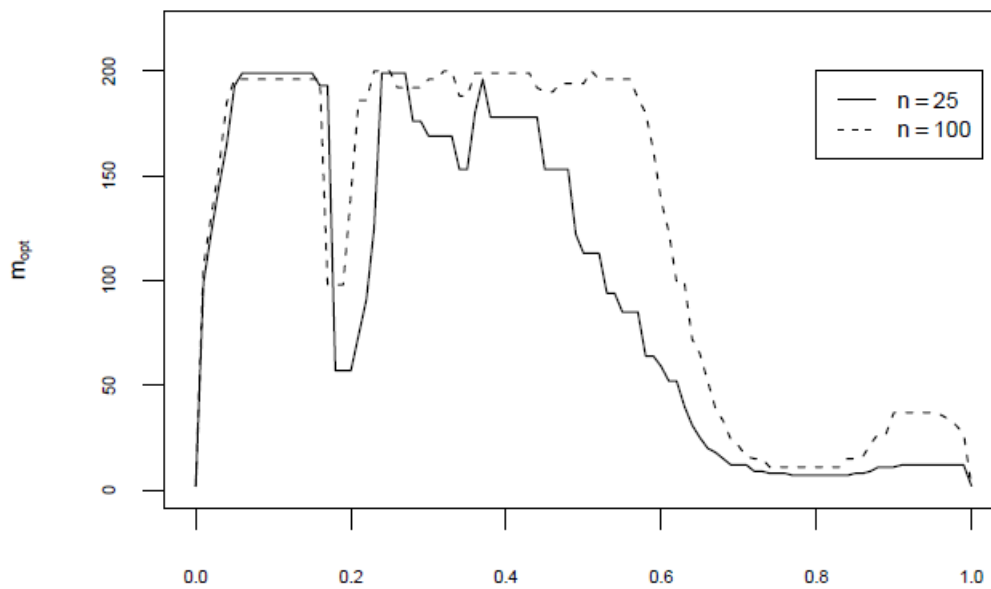


FIGURE 4.9. Valeur optimale de m pour un échantillon de loi $1/2\text{Beta}(12,6)+1/2\text{Beta}(9,1)$.

4.2. DENSITÉ

Nous avons considéré ici 3 distributions : (i) $\text{beta}(2,5)$, (ii) $\chi^2(2)$ et (iii) $t(2)$. Les estimateurs de Bernstein n'étant définis que sur l'intervalle I pour les illustrations, suivant le travail de Babu et al. (2002) nous avons utilisé les transformations suivantes $Y = X/(1 + X)$ pour le cas (ii) et $Y = 1/2 + (1/\pi) \arctan(X)$ pour le cas (iii). D'après le théorème 3.11, le choix $m = o(n/\log n)$ est suffisant pour garantir une bonne estimation de f à partir de $\tilde{f}_{n,m}$. Nous avons fait les choix $m = n/\log n$ et $m = n$. Nous allons comparer l'estimateur de Bernstein à l'estimateur par noyaux qui est défini par

$$\hat{f}_{n,h}(x) = \frac{1}{nh} \sum_{k=1}^n K\left(\frac{x - X_k}{h}\right)$$

où K est le noyau et h la fenêtre de l'estimateur. Suivant les recommandations de Silverman (1986) sur le paramètre h , nous avons utilisé la fonction `density` de R avec le noyau gaussien et le choix par défaut du paramètre de lissage. Les figures 4.13, 4.14 et 4.15 nous montrent le comportement des estimateurs pour 10 simulations de taille $n = 75$ et $n = 110$. On remarque que globalement l'estimateur de Bernstein avec $m = n/\log n$ est moins variable comparé au cas $m = n$ et à l'estimateur par noyau. On note cependant que l'estimateur de Bernstein présente des erreurs d'estimation aux bornes, qui sont plus visibles dans le cas (i). Le travail récent de Leblanc (2012b) porte sur ce sujet.

Pour ces simulations, nous avons fait le choix d'utiliser le noyau gaussien. Il est possible qu'un autre choix de noyau ou du paramètre de lissage nous donne des résultats contraires à ceux présentés ici. De même le choix de la transformation pour les fonctions de densité dont le support est différent de I pourrait avoir un effet non négligeable sur l'estimation.

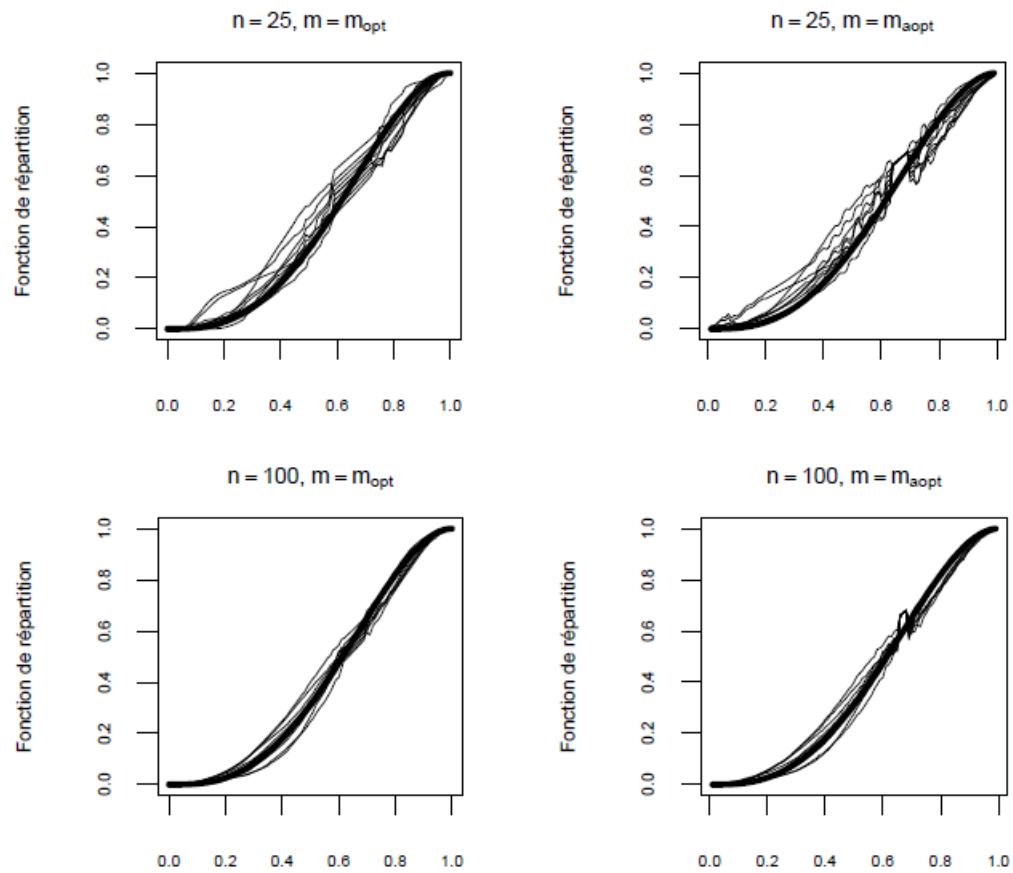


FIGURE 4.10. Estimateur de Bernstein pour un échantillon de loi Beta(3,2) : cas m optimal.

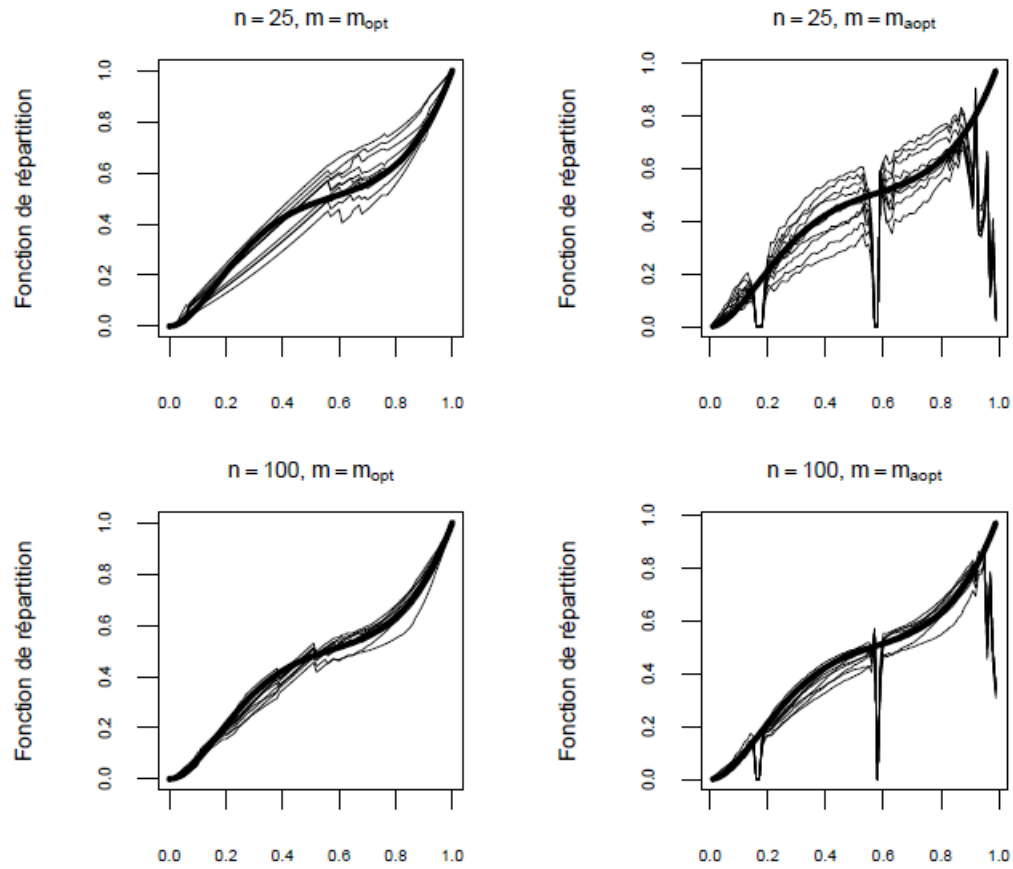


FIGURE 4.11. Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$: cas m optimal.

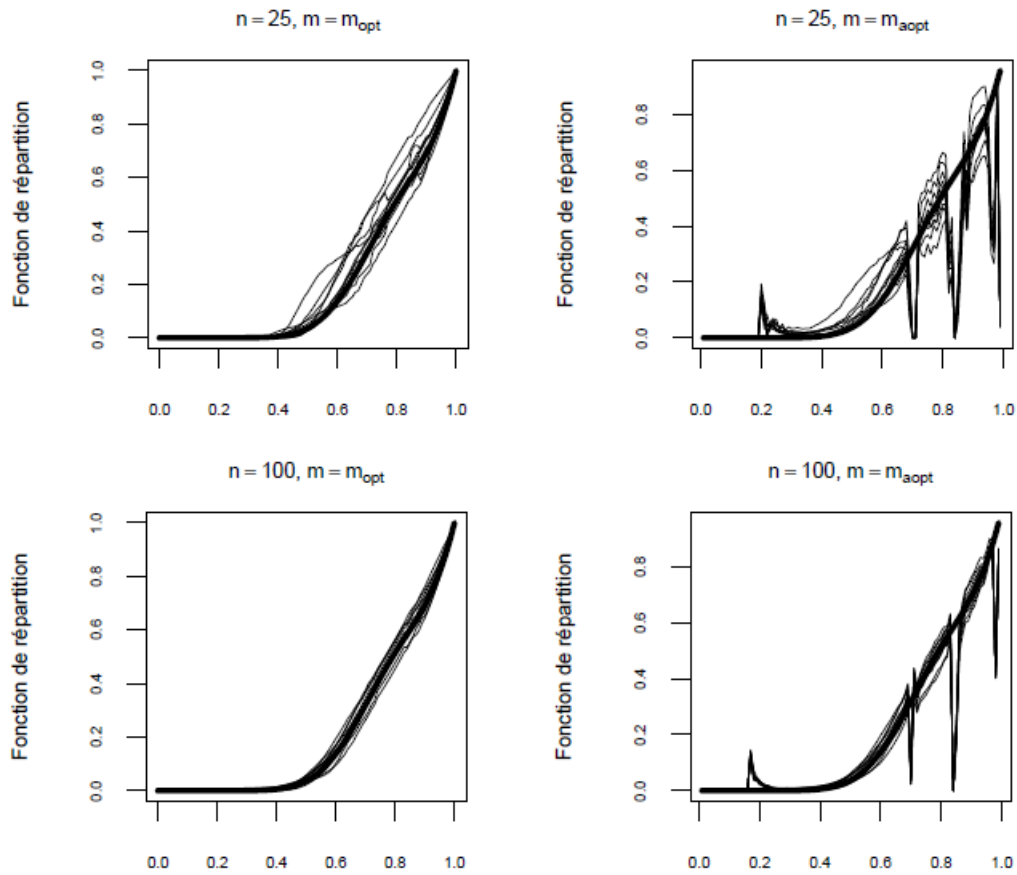


FIGURE 4.12. Estimateur de Bernstein pour un échantillon de loi $1/2\text{Beta}(2,6)+1/2\text{Beta}(6,1)$: cas m optimal.

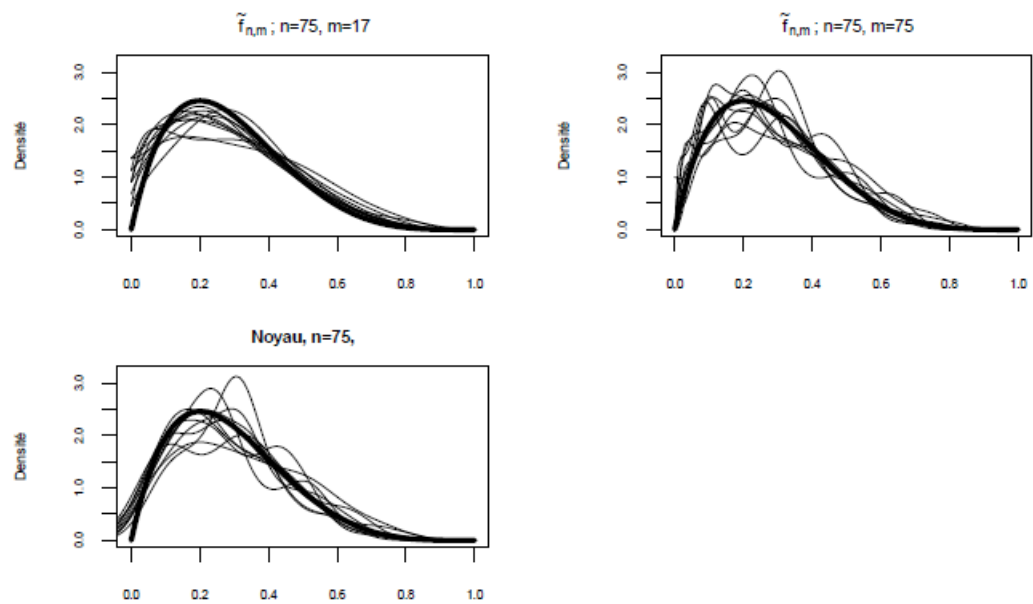
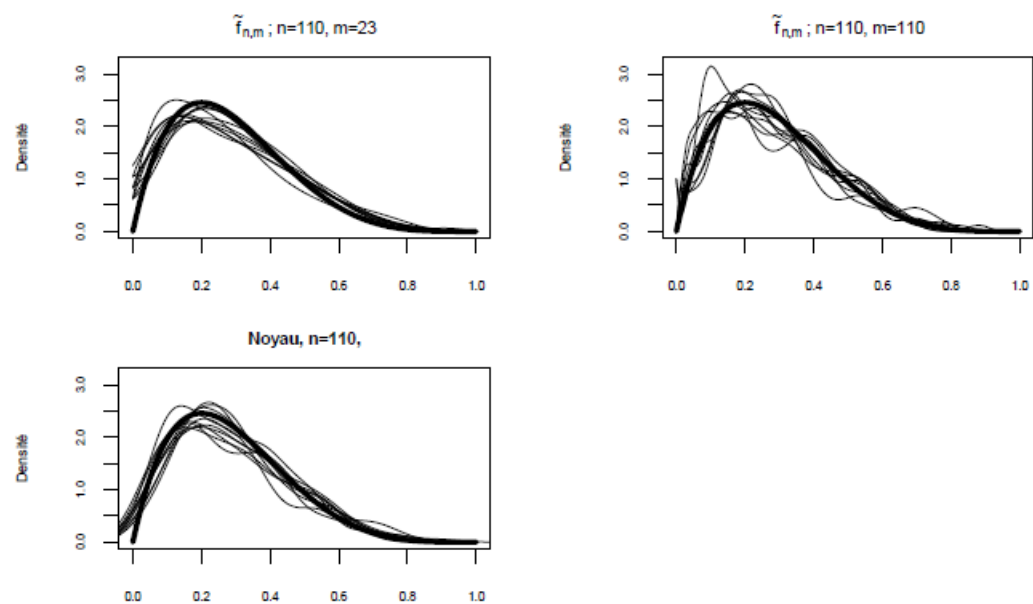
(a) $n = 75$ (b) $n = 110$

FIGURE 4.13. Estimateur de Bernstein pour un échantillon de loi Beta(2,5).

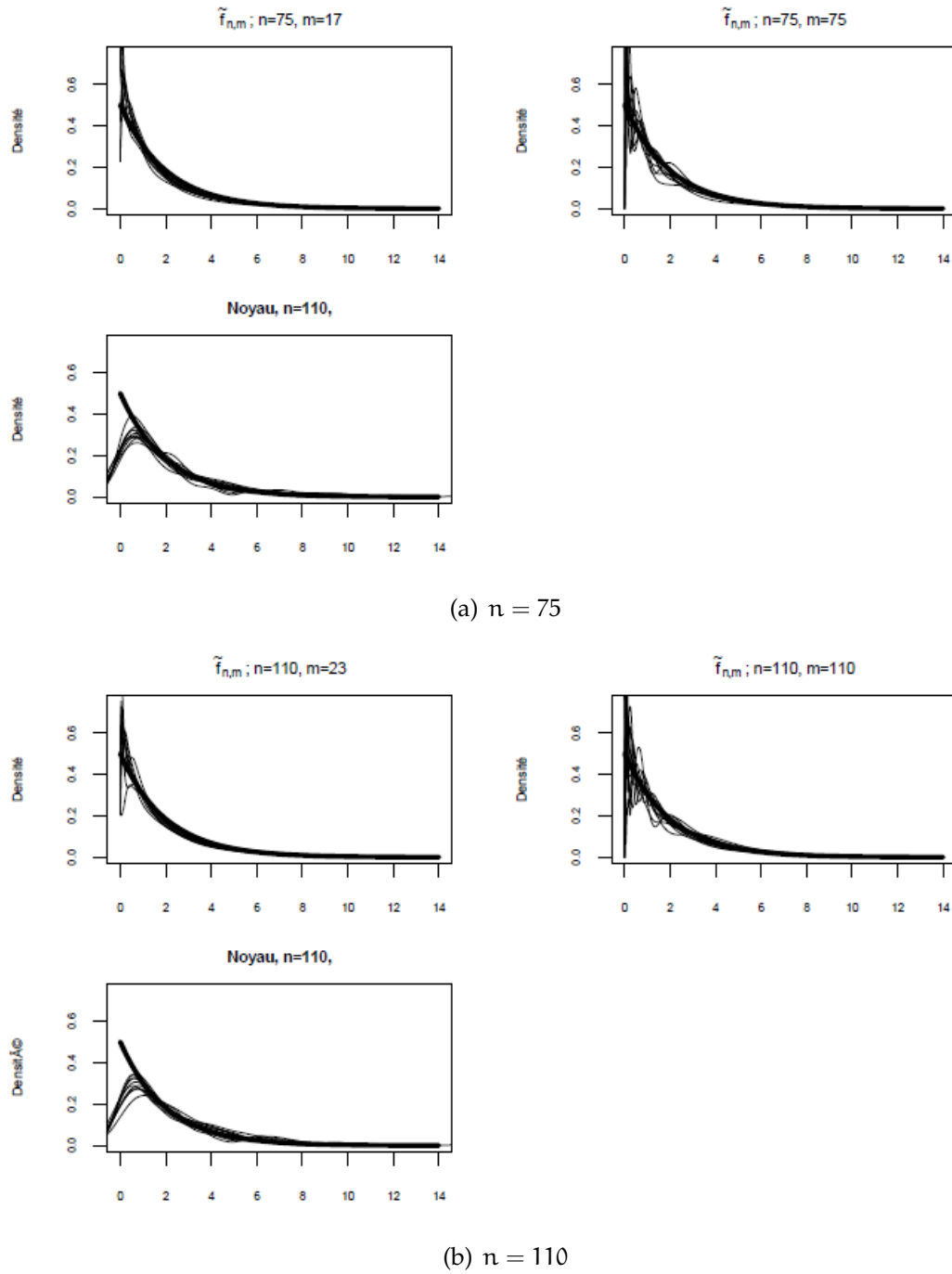
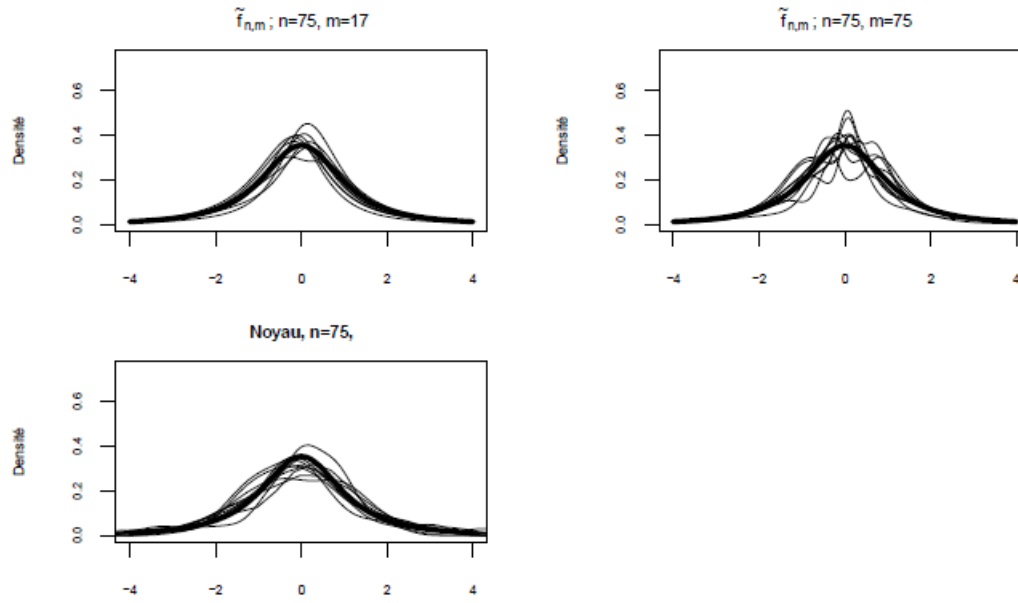
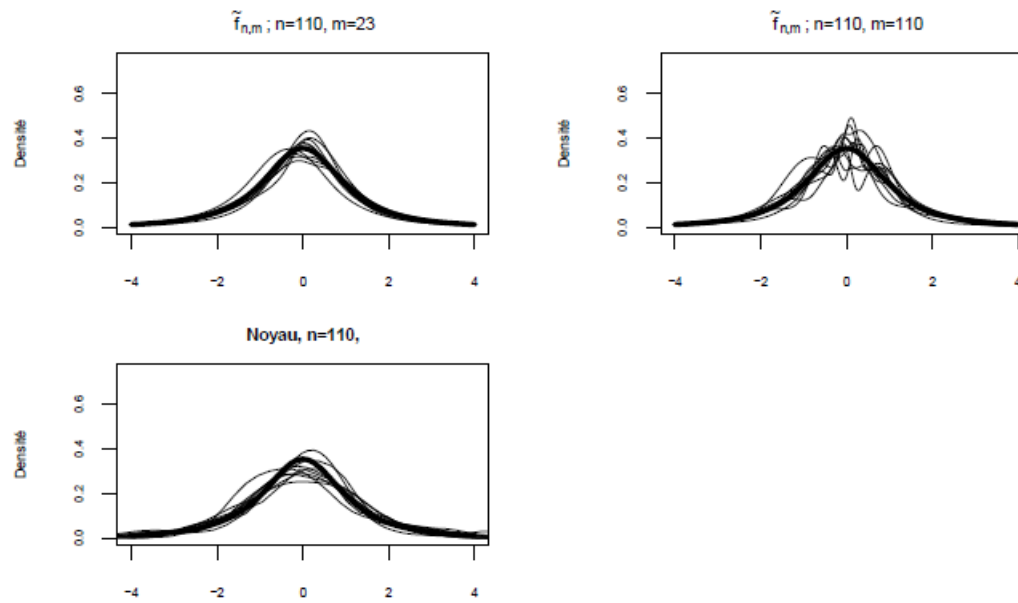


FIGURE 4.14. Estimateur de Bernstein pour un échantillon de loi $\chi^2(2)$.

(a) $n = 75$ (b) $n = 110$ FIGURE 4.15. Estimateur de Bernstein pour un échantillon de loi $t(2)$.

CONCLUSION

Le but de ce travail était de présenter une alternative récente aux différents estimateurs de fonction de répartition et de densité classiques. Les estimateurs de Bernstein sont des polynômes dont les coefficients dépendent de l'échantillon d'observations prélevées, ce qui les rend particulièrement maniables et faciles à mettre en oeuvre à partir d'un logiciel. Nous nous sommes intéressés dans un premier temps à l'origine des polynômes de Bernstein. Après les avoir définis nous avons montré leurs principales propriétés, parmi lesquelles le théorème de Weierstrass et les résultats avec le module de continuité. Nous avons ensuite montré d'importantes inégalités concernant le module de continuité du processus empirique dans un cadre multivarié. Nous avons enfin introduit les estimateurs de Bernstein. Nous avons conclu que l'estimateur de Bernstein pour la fonction de répartition était, sous des conditions de dérivabilité, asymptotiquement sans biais et convergent en moyenne quadratique. Nous avons aussi montré que les estimateurs de Bernstein étaient asymptotiquement normaux et possédaient comme les estimateurs classiques, la propriété de Chung-Smirnov. De plus, pour une taille d'échantillon fixée, il a été possible de trouver un estimateur de Bernstein plus efficient en moyenne quadratique que la fonction de répartition empirique.

Enfin, quelques simulations nous ont permis d'apprécier le comportement des estimateurs de Bernstein. Nous avons pu remarquer que l'estimateur de Bernstein se comportait mieux en présence de faible courbure de la fonction de répartition. Nous avons aussi remarqué que le choix du m optimal était problématique dans les situations où l'EQM ne varie pas beaucoup pour un point de discrétisation fixé.

Finalement, nous avons conclu que le choix asymptotique de m fournissait une estimation acceptable de la fonction de répartition en dehors des voisinages des points critiques de la fonction de densité.

BIBLIOGRAPHIE

- Babu, G.J., A.J. Canty & Y.P. Chaubey (2002), 'Application of Bernstein Polynomials for smooth estimation of a distribution and density function', *Journal. Statist. Plan. Inf.* **105**, 377–392.
- Bernstein, S. N. (1912), 'Démonstration du théorème de weierstrass fondée sur le calcul des probabilités', *Commun. Soc. Math. Kharkow* **13**, 1–2.
- Cressie, N. (1978), 'A finely tuned continuity correction', *Annals of the Inst. of Stat. Math.* **30**, 435–442.
- Einmahl, J.H.J. (1987), Multivariate Empirical Processes, Technical report, CWI Tract 32, Amsterdam.
- Einmahl, J.H.J. & F.H. Ruymgaart (1987), 'The almost sure behavior of the oscillation modulus of the multivariate empirical process', *Stat. Probab. Lett.* **6**, 87–96.
- Feller, W. (1965), *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, New York.
- Jones, O., R. Maillardet & A. Robinson (2009), *Introduction to Scientific Programming and Simulation Using R*, Chapman & Hall/CRC, Boca Raton, FL.
- Leblanc, A. (2009), 'Chung-Smirnov property for Bernstein estimators of distribution functions', *Journal of Nonparametric Stat.* **21**, 133–142.
- Leblanc, A. (2012a), 'On estimating distribution functions using Bernstein polynomials', *Annals of the Inst. of Stat. Math.* **64**, 919–943.
- Leblanc, A. (2012b), 'On the boundary properties of Bernstein polynomial estimators of density and distribution functions', *Journal. Statist. Plan. Inf.* **142**, 2762–2778.
- Lorentz, G.G. (1986), *Bernstein Polynomials*, 2nd ed., Chelsea Publishing, New York.
- Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

- Shorack, G.R. & J.A. Wellner (1986), *Empirical Processes with Applications to Statistics*, John Wiley & Sons, New York.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
- Vitale, R.A. (1975), 'A Bernstein polynomial approach to density estimation', *Statistical Inference and Related Topics* **2**, 87–100.
- Winter, B.B. (1979), 'Convergence rate of perturbed empirical distribution functions', *J.Appl. Probab* **16**, 163–173.

Annexe A

RÉSULTATS IMPORTANTS

A.1. THÉORÈME DE LINDEBERG

Définition A.1. Soit X une variable aléatoire. Elle est dite de type 'treillis' si elle prend les valeurs de la forme : $b, b \pm h, b \pm 2h, \dots$ avec $b \geq 0$ et $h > 0$. La constante h est appelée le pas de la distribution.

Théorème A.1. Soit X_1, X_2, \dots des variables aléatoires i.i.d de même loi F de type 'treillis', vérifiant $E(X_i) = 0$ et $E(X_i^2) = 1$. On pose $S_n = X_1 + X_2 + \dots + X_n$ et $x = (nb + kh)/\sqrt{n}$ (b et h comme dans la Définition A.1) où $k = 0, \pm 1, \pm 2, \dots$. On a lorsque $n \rightarrow \infty$:

$$\frac{\sqrt{n}}{h} P\left(\frac{S_n}{\sqrt{n}} = x\right) \rightarrow \varphi(x), \quad (\text{A.1})$$

où φ est la densité de la loi normale centré réduite.

Preuve. La preuve est celle du théorème 3, section VI.5 (P. 490) de Feller (1965). \square

Nous énonçons maintenant le théorème de Lindeberg,

Théorème A.2. Soit X_1, X_2, \dots des variables aléatoires indépendantes, vérifiant $E(X_k) = 0$ et $E(X_k^2) = \sigma_k^2$, pour tout k . On pose $S_n = X_1 + X_2 + \dots + X_n$ et $s_n^2 = \text{Var}(S_n) = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$. Si pour tout ε positif

$$\frac{1}{s_n^2} \sum_{k=1}^n E\left(X_k^2 \mathbb{1}_{\{|X_k| \geq \varepsilon s_n\}}\right) \rightarrow 0, \quad n \rightarrow \infty, \quad (\text{A.2})$$

alors

$$S_n/s_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1). \quad (\text{A.3})$$

A-ii

Preuve. La preuve est celle du théorème 1, section VI.6 (P. 491-492) de Feller (1965). □

A.2. THÉORÈME DE DINI

Nous n'énonçons que le premier théorème de Dini.

Théorème A.3. *Soit $(u_n)_{n \in \mathbb{N}}$ une suite de fonctions continues sur I à valeurs dans \mathbb{R} . On suppose que u_n converge simplement vers une fonction continue u . On a le résultat suivant : si chaque u_n est croissante alors la convergence est uniforme.*

Preuve. Soit $\varepsilon > 0$, puisque u est continue sur I elle est bornée et uniformément continue. Posons $\eta = \omega_u(\varepsilon)$. Considérons $a_0 = 0 < a_1 < \dots < a_{m-1} < a_m = 1$ une subdivision de I telle que $|a_i - a_{i-1}| \leq \eta$, pour $i = 1, 2, \dots, m$. Il existe $n_0 \in \mathbb{N}$ tel que $n \geq n_0$ implique $|u(a_i) - u_n(a_i)| \leq \varepsilon$, $i = 0, 1, \dots, m$. Soit maintenant $x \in I$, il existe $i \in \{0, 1, \dots, m-1\}$ tel que $a_i \leq x \leq a_{i+1}$, par continuité on a $|u(x) - u(a_i)| \leq \varepsilon$. Pour $n \geq n_0$ on peut alors écrire en utilisant la croissance de u_n que

$$\begin{aligned} |u(x) - u_n(x)| &= |u(x) - u(a_i)| + |u(a_i) - u_n(a_i)| + |u_n(a_i) - u_n(x)|, \\ &\leq 2\varepsilon + u_n(x) - u_n(a_i), \\ &\leq 2\varepsilon + u_n(a_{i+1}) - u_n(a_i), \\ &\leq 2\varepsilon + |u_n(a_{i+1}) - u(a_{i+1})| + |u(a_{i+1}) - u_n(a_i)| + |u(a_i) - u_n(a_i)|, \\ &\leq 5\varepsilon \end{aligned}$$

La convergence est donc uniforme. □

Annexe B

CODES R UTILISÉS POUR LES SIMULATIONS

Nous ne présentons que les cas (iii) des sections 4.1 et 4.2. Les autres cas s'obtiennent de manière similaire.

B.1. CAS (III) SECTION 4.1

```
#La fonction suivante permet de calculer  
#l'estimateur de Bernstein pour la fonction de répartition.
```

```
F.n.m=function(x,y,m){  
  #y est le vecteur de réalisation de la loi  
  #On veut évaluer F.n.m en x  
  p=length(x)  
  z=rep(0,p)  
  u=0:m/m  
  for (i in 1:p){  
    z[i]=sum((ecdf(y)(u))*dbinom(0:m,m,x[i]))}  
  return(z)}  

```

```
#### Simulation pour la  $1/2 \times \text{Beta}(12,6) + 1/2 \times \text{Beta}(9,1)$  ###  
# Cas n=25#
```

```
n= 25  
set.seed(31)
```

B-ii

```
v=rmultinom(n*10,1,c(1/2,1/2))
u=matrix(c(rbeta(n*10,12,6),rbeta(n*10,9,1)),nrow=2,byrow=T)
data=apply(u*v,2,sum)
data=matrix(data,ncol=10,nrow=n)
z=seq(0,1,0.01)
```

```
m=8
par(mfrow=c(1,2),mar=c(2.5,4,2,2), cex.main=.7,
    cex.axis=0.5,cex.lab=.65)
plot(z,1/2*pbeta(z,12,6)+1/2*pbeta(z,9,1),type='l',
     lwd=3,main="n=25, m=8",ylab="Fonction de répartition")
for (i in 1:10){
  lines(z,F.n.m(z,data[,i],m),type="l",lwd=0.4) }
```

```
m=25
plot(z,1/2*pbeta(z,12,6)+1/2*pbeta(z,9,1),type='l',lwd=3,
     main="n=25, m=25",ylab="Fonction de répartition")
for (i in 1:10){
  lines(z,F.n.m(z,data[,i],m),type="l",lwd=0.4) }
```

```
#### les graphiques pour les cas n=75 et
#### 100 s'obtiennent de manière similaire.
```

```
#####Cacul de l'EQM #####
##n=25##
```

```
n=25
M=1000
z=seq(0,1,0.01)
set.seed(12)
v=rmultinom(n*M,1,c(1/2,1/2))
```

```

u=matrix(c(rbeta(n*M,12,6),rbeta(n*M,9,1)),nrow=2,byrow=T)
data=apply(u*v,2,sum)
data=matrix(data,ncol=M,nrow=n)
m=c(8,25)
A1=matrix(0,ncol=101,nrow=M)
A2=matrix(0,ncol=101,nrow=M)
A3=matrix(0,ncol=101,nrow=M)
for (i in 1:M){
A1[i,]= (F.n.m(z,data[,i],m[1])-(pbeta(z,12,6)+pbeta(z,9,1))/2)^2
A2[i,]= (F.n.m(z,data[,i],m[2])-(pbeta(z,12,6)+pbeta(z,9,1))/2)^2
A3[i,]= (ecdf(data[,i])(z)-(pbeta(z,12,6)+pbeta(z,9,1))/2)^2}

##Calcul du m optimal##
B1=matrix(0,ncol=101,nrow=199)
z=seq(0,1,0.01)
mopt=rep(0,101)
for (m in 2:200){
A4=matrix(0,ncol=101,nrow=M)
for (i in 1:M){
A4[i,]= (F.n.m(z,data[,i],m)-(pbeta(z,12,6)+pbeta(z,9,1))/2)^2
}
B1[m-1,]=apply(A4,2,mean)}
mopt_25=apply(B1,2,which.min)+1
bb4=apply(B1,2,min)
bb1=apply(A1,2,mean)
bb2=apply(A2,2,mean)
bb3=apply(A3,2,mean)
par(mfrow=c(2,2),mar=c(2.5,4,2,2), cex.main=.8,
cex.axis=0.8,cex.lab=.8)
plot(z,bb3*1000,type="l",ylab=expression(paste("EQM " "%*% 10^-3 ")),
main="n=25",lty=4,ylim=c(0,12))

```

B-iv

```
lines(z,bb2*1000,type="l",lty=2)
lines(z,bb1*1000,type="l",lty=1)
s3 <- expression(F[n], m==8,m==25 m[opt])
utils::str(legend(0.02,12,s3, lty=c(4,1,2), plot=T,
adj = c(0, .5),lwd=0.2,title.adj = 0.2, seg.len = 1,cex=.7))

#### les simulations pour le cas n=100
#### s'obtiennent de manière similaire.

par(mfrow=c(1,1),mar=c(2.5,4,2,2), cex.main=.7,
cex.axis=0.5,cex.lab=.65)
plot(z,mopt_25, ylab="", xlab=expression(m[opt]))
lines(z,mopt_100, ylab="", xlab=expression(m[opt]))
s3 <- expression(n==25,n==100)
utils::str(legend(0.5,140,s3, lty=c(1,2), plot=T,
adj = c(0, .5),lwd=0.8,title.adj = 0.2, seg.len = 1.7,cex=.7))
```

B.2. CAS (III) SECTION 4.2

```
#La fonction suivante permet de calculer
#l'estimateur de Bernstein pour la fonction de densité.

f.n.m=function(x,y,m){
#y est le vecteur de réalisation de la loi
#On veut évaluer f.n.m en x
p=length(x)
z=rep(0,p)
u=1:m/m
v=0:(m-1)/m
for (i in 1:p){
z[i]=m*sum((ecdf(y)(u)-ecdf(y)(v))*dbinom(0:(m-1),m-1,x[i]))}
return(z)}
```

```

####Loi Khi^2(2) ####

## Cas n=75#
n=75
set.seed(000)
data=rchisq(n*10,df=2)
data=matrix(data,ncol=10,nrow=n)
z1=seq(0,14,0.01)
z=z1/(1+z1)

m=17
par(mfrow=c(2,2),mar=c(2.5,4,2,2), cex.main=.7,
    cex.axis=0.5,cex.lab=.65)
plot(z1,dchisq(z1,df=2),type='l',ylim=c(0,0.75),lwd=3,
     main=expression(paste(tilde(f)[paste(n,"m")], " ; n=75", " , m=17")),
     ylab="Densité")
for (i in 1:10){
  lines(z1,1/(1+z1)^2*f.n.m(z,data1[,i],m),type="l",lwd=0.4) }

m=75
plot(z1,dchisq(z1,df=2),ylim=c(0,0.75),type='l',lwd=3,
     main=expression(paste(tilde(f)[paste(n,"m")], " ; n=75", " , m=75")),
     ylab="Densité")
for (i in 1:10){
  lines(z1,1/(1+z1)^2*f.n.m(z,data1[,i],m),type="l",lwd=0.4) }

plot(z1,dchisq(z1,df=2),type='l',ylim=c(0,0.75),lwd=3,main="Noyau,
     n=75",ylab="Densité")
for (i in 1:10){
  lines(density(data[,i], bw = "nrd0", adjust = 1,weights = NULL,

```

B-vi

```
kernel = "gaussian",  
n = 101, 0, 1, cut = 3),type="l",lwd=0.4)}
```

##Les simulations dans le cas n=110 s'obtiennent de manière similaire.